

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 15:05:46 ; Search time 12084 Seconds
(without alignments)
11373.780 Million cell updates/sec

Title: US-09-721-114-1_COPY_655_3825

Perfect score: 3171

Sequence: 1 atggagattgttcagtaga.....acgagtattggtaccgc 3171

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.scs.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vit.*

38: em.sv.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3171	100.0	4310	6	BD083699	BD083699 Novel gen
2	3167.8	99.9	3896	8	AF326768	AF326768 Oryza sat
3	2978.4	93.9	3626	8	AK103364	AK103364 Oryza sat
4	2573.2	81.1	9455	6	BD083700	BD083700 Novel gen
5	2573.2	81.1	150594	8	AP001859	AP001859 Oryza sat
6	62.8	2.0	7218	6	I66494	I66494 Sequence 14
7	62	2.0	2000	6	AX655393	AX655393 Sequence
8	52	1.6	2000	6	AX655393	AX655393 Sequence
9	50.8	1.6	180023	2	EX897685	EX897685 Danio rer
10	49.8	1.6	178463	2	EX248504	EX248504 Danio rer
11	48.8	1.5	185533	2	EX005453	EX005453 Danio rer
12	48.6	1.5	228730	2	AC096088	AC096088 Rattus no
13	48.4	1.5	124927	5	EX510987	EX510987 Zebrafish
14	48.4	1.5	250029	3	AB014820	AB014820 Plasmodiu
15	48.2	1.5	141041	2	AC026135	AC026135 Homo sapi
16	48.2	1.5	148130	9	AC092925	AC092925 Homo sapi
17	47.8	1.5	213544	2	EX510939	EX510939 Danio rer
18	47.2	1.5	106172	9	AL591122	AL591122 Human DNA
19	46.8	1.5	225453	2	AC097957	AC097957 Rattus no
20	46.6	1.5	127494	2	AC117775	AC117775 Mus muscu
21	46.6	1.5	187418	10	AL672278	AL672278 Mouse DNA
22	46.2	1.5	155304	9	AC055845	AC055845 Homo sapi
23	46.2	1.5	167446	2	AC137718	AC137718 Homo sapi
24	46.2	1.5	189203	2	AC026077	AC026077 Homo sapi
25	45.6	1.4	198743	2	EX530070	EX530070 Danio rer
26	45.4	1.4	274	6	AR248600	AR248600 Sequence
27	45.4	1.4	145120	8	AP003434	AP003434 Oryza sat
28	45.2	1.4	1141	6	AX083744	AX083744 Sequence
29	45.2	1.4	93100	9	AL162614	AL162614 Human DNA
30	45.2	1.4	94534	5	AL929250	AL929250 Zebrafish
31	45.2	1.4	114980	2	AC087825	AC087825 Homo sapi
32	45.2	1.4	158392	2	AC022048	AC022048 Homo sapi
33	45	1.4	173233	9	AC116002	AC116002 Homo sapi
34	45	1.4	217726	10	AL928680	AL928680 Mouse DNA
35	45	1.4	330050	3	PPA929355	PPA929355 Plasmodiu
36	44.8	1.4	139074	2	AC131191	AC131191 Mus muscu
37	44.8	1.4	165488	10	AL807810	AL807810 Mouse DNA
38	44.4	1.4	3400	14	SH10145A	L20708 Pseudorabie
39	44.4	1.4	5200	14	SHE58096S	AJ580965 Suid herp
40	44.4	1.4	151802	3	AC114263	AC114263 Dictyoste
41	44.4	1.4	246642	2	AC123079	AC123079 Rattus no
42	44.2	1.4	95097	9	AL133232	AL133232 Human DNA
43	44.2	1.4	177747	10	AC122051	AC122051 Mus muscu
44	44.2	1.4	182528	2	AC126422	AC126422 Mus muscu
45	44.2	1.4	195509	2	AC107868	AC107868 Mus muscu

ALIGNMENTS

RESULT 1
BD083699
LOCUS
DEFINITION Novel gene participating in response to brassinosteroid.
ACCESSION BD083699
VERSION BD083699.1 GI:22629309
KEYWORDS JP 2001327287-A/1.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 4310)
Hirochika, H., Yamazaki, T. and Miyao, A.
REFERENCE
AUTHORS
TITLE Novel gene participating in response to brassinosteroid

JOURNAL Patent: JP 2001327287-A 1 27-NOV-2001;
 NATL INST OF AGROBIOLOGICAL RESOURCES, BIO ORIENTED TECHNOLOGY
 RESEARCH ADVA JAPAN FORAGE SEED ASSOCIATION
 COMMENT OS *Oryza sativa* (rice)
 PN JP 2001327287-A/1
 PD 27-NOV-2001
 PF 19-MAY-2000 JP 2000149106
 PI HIROHIKO HIROCHIKA, TOKIO YAMAZAKI, AKIO MIYAO
 PC C12N15/09, C12N15/00
 CC Novel gene participating in response to brassinosteroid FH
 key Location/Qualifiers
 FT CDS (635)..(3828).
 FEATURES
 source Location/Qualifiers
 1..4310
 /organism="Oryza sativa"
 /mol_type="genomic DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Query Match 100.0%; Score 3171; DB 6; Length 4310;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCGAATTGTTGCAGTAGATCAGGAGGAGCTCGTGTGTTTGGACGCAACTGTATGCTT 60
 DB 655 ATGCGAGTTGTTGCAGTAGATCAGGAGGAGCTCGTGTGTTTGGACGCAACTGTATGCTT 714
 QY 61 GTCGTGTGTGAACTCGGTGCTGTAGCGCCAGTTGTGGAGCTGACGCGGCTGCTCAG 120
 DB 715 GCTCGTGTGTGAACTCGGTGCTGTAGCGCCAGTTGTGGAGCTGACGCGGCTGCTCAG 774
 QY 121 GATGCAAGCCGCTGAAGCTGTGTAGACGAACCGGCACAAACCAATTCGAGCAATTTCTCC 180
 DB 775 GATGCAAGCCGCTGAAGCTGTGTAGACGAACCGGCACAAACCAATTCGAGCAATTTCTCC 834
 QY 181 ATAAGAGGTATGTTGCTCTTCTTCTAGAGAAGATCCAAATTCGCTCTCTATCTCGG 240
 DB 835 ATAAGAGGTATGTTGCTCTTCTTCTAGAGAAGATCCAAATTCGCTCTCTATCTCGG 894
 QY 241 ATTTCATGACACAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCAATTTCTGTA 300
 DB 895 ATTTCATGACACAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCAATTTCTGTA 954
 QY 301 GCAAGTTTCGACATGGGATTTGCTCGAAGTGGTGGATGATGTAAGTTGAAACTTCAGATAAT 360
 DB 955 GCAAGTTTCGACATGGGATTTGCTCGAAGTGGTGGATGATGTAAGTTGAAACTTCAGATAAT 1014
 QY 361 GGAACAGCACCAAGAACTCTTCCGCAAAACAGATGGCAAAAGTGTATGTTGCTCCATC 420
 DB 1015 GGAACAGCACCAAGAACTCTTCCGCAAAACAGATGGCAAAAGTGTATGTTGCTCCATC 1074
 QY 421 ACATTTGTCGGAGCACTTTTGGCCCTGTAGTGTGGTTCCTCCAAAGAGTCTCCCTAGC 480
 DB 1075 ACATTTGTCGGAGCACTTTTGGCCCTGTAGTGTGGTTCCTCCAAAGAGTCTCCCTAGC 1134
 QY 481 ACACATATCTCTCAAGGAAGAATGCTGTATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 540
 DB 1135 ACACATATCTCTCAAGGAAGAATGCTGTATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194
 QY 541 GGCATGACTCCAAATGCAATGCGCCTTCTGGCAGGAATGAGCTGCTGAGGCCAATACT 600
 DB 1195 GGCATGACTCCAAATGCAATGCGCCTTCTGGCAGGAATGAGCTGCTGAGGCCAATACT 1254
 QY 601 GATTACCAATGAAAGATTTGCAAGGCCAGCCCAAAATTTATGATGGCAGCAAAATGTC 660
 DB 1255 GATTACCAATGAAAGATTTGCAAGGCCAGCCCAAAATTTATGATGGCAGCAAAATGTC 1314
 QY 661 TCTGAGGACAAACACTTCTGTGTGATGTTGGGGCTTTACTGTAAGTTCCCGAGATTACATGG 720
 DB 1315 TCTGAGGACAAACACTTCTGTGTGATGTTGGGGCTTTACTGTAAGTTCCCGAGATTACATGG 1374
 QY 721 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAACATTTCTGAAGTGGTC 780

Query Match	99.9%;	Score 3167.8;	DB 8;	Length 3896;
Best local Similarity	99.9%;	Pred. No. 0;		
Matches 3169;	Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;
QY	1	ATGGAGATTGTTCCAGTAGATCAGGAGGAGCTCGTGTGTTGTTGGGACGAACTGTATGCTT	60	
DB	241	ATGGAGATTGTTCCAGTAGATCAGGAGGAGCTCGTGTGTTGTTGGGACGAACTGTATGCTT	300	
QY	61	GCTCGTGTGGAACTGGTGTCTGTAGCCGACGTCTGGAGCTGACAGCGACGCTCGTCAG	120	
DB	301	GCTCGTGTGGAACTGGTGTCTGTAGCCGACGTCTGGAGCTGACAGCGACGCTCGTCAG	360	
QY	121	GATCGACCCGCTGAAGCTGGTGTAGACGAAACCGGCAACAACCAATGCGAGCAATTTCTCC	180	
DB	361	GATCGACCCGCTGAAGCTGGTGTAGACGAAACCGGCAACAACCAATGCGAGCAATTTCTCC	420	
QY	181	ATAGAGGGTATGTTGCTCTCTTTCAGAGCAAGATCCAAAATTTCTGCTCTATCTCGG	240	
DB	421	ATAGAGGGTATGTTGCTCTCTTTCAGAGCAAGATCCAAAATTTCTGCTCTATCTCGG	480	
QY	241	ATTTTCATGACACGAAAAAATGTATGATGAACAACAAGCTAGTTCAAGCCCATTTCTGTA	300	
DB	481	ATTTTCATGACACGAAAAAATGTATGATGAACAACAAGCTAGTTCAAGCCCATTTCTGTA	540	
QY	301	GCAAGTTTCAGAGTCGGATGCTCGAAGTCTTGGATGATGTTGAAACTTCAGATAAT	360	
DB	541	GCAAGTTTCAGAGTCGGATGCTCGAAGTCTTGGATGATGTTGAAACTTCAGATAAT	600	
QY	361	GGAAACAGCACCAGAAACTCTTCCCGCAAAAGCAAGTGGCAAGTGTGCTCCATC	420	
DB	601	GGAAACAGCACCAGAAACTCTTCCCGCAAAAGCAAGTGGCAAGTGTGCTCCATC	660	
QY	421	ACATTTGTTCCGAGCATTTTGTGCGCTGTAGTGTGTTGCCAAAAGTGTCTCTAGC	480	
DB	661	ACATTTGTTCCGAGCATTTTGTGCGCTGTAGTGTGTTGCCAAAAGTGTCTCTAGC	720	
QY	481	ACACAATCATCTCAAGGGAAGAAATGCTGATAGATCAAACTCTTCCAAAGAGTGTGCAAGAA	540	
DB	721	ACACAATCATCTCAAGGGAAGAAATGCTGATAGATCAAACTCTTCCAAAGAGTGTGCAAGAA	780	
QY	541	GGCAATGACTCCAAATGCAATGCAGCTCTTCTGGCAAGAAATGAGCTGCTGAGGCCCAATACT	600	
DB	781	GGCAATGACTCCAAATGCAATGCAGCTCTTCTGGCAAGAAATGAGCTGCTGAGGCCCAATACT	840	
QY	601	GATTCACCAATGAAGATTTCCAGGGCCAGCCCAAAATATGATGTGGCAGCAAAATGTC	660	
DB	841	GATTCACCAATGAAGATTTCCAGGGCCAGCCCAAAATATGATGTGGCAGCAAAATGTC	900	
QY	661	TCTGAGGACACACTTCGTTTGATGTTGGGCTTTTACTCTGAAGTTTCCCAAGATTACATGG	720	
DB	901	TCTGAGGACACACTTCGTTTGATGTTGGGCTTTTACTCTGAAGTTTCCCAAGATTACATGG	960	
QY	721	CACATAGAAATTAATGTTGTCAGATCAAACTCATCCACTCCAAACTTTCTGAAGTGGTC	780	
DB	961	CACATAGAAATTAATGTTGTCAGATCAAACTCATCCACTCCAAACTTTCTGAAGTGGTC	1020	
QY	781	CTCAAAAGAAATGAAGATGAAAAATGAAAAAACTGAAGAGACTCTTGTGCTGACAGTGC	840	
DB	1021	CTCAAAAGAAATGAAGATGAAAAATGAAAAAACTGAAGAGACTCTTGTGCTGACAGTGC	1080	
QY	841	AAATTTGACCAAGATTCCTAAACCAATGTTCTGAAAGGAACTGATCAGTGTGCTGAGCAG	900	
DB	1081	AAATTTGACCAAGATTCCTAAACCAATGTTCTGAAAGGAACTGATCAGTGTGCTGAGCAG	1140	
QY	901	TGCAATTTGACCAAGATTCGAAACCAAGTGTCTGGGACAGAAATGTGAGCAGATCTGCAAT	960	
DB	1141	TGCAATTTGACCAAGATTCGAAACCAAGTGTCTGGGACAGAAATGTGAGCAGATCTGCAAT	1200	
QY	961	GAGCCATGTGAAGAAGTTGTTCTCAAAAAGAGCTCCAAATCTAAGAGGAGACCGATTAAG	1020	
DB	1201	GAGCCATGTGAAGAAGTTGTTCTCAAAAAGAGCTCCAAATCTAAGAGGAGACCGATTAAG	1260	

[illegible]

2341 CCATCTCTCAGAACTTTCACTTACTCAGCAACAGCAGACACATTTGGCGATGGAAGAA 2400
2161 ATGCTCACTATTGCTGCAAGCTCACCATTCTTTTCACTCATGATGATCATGATATTGCT 2220
2401 ATGCTCACTATTGCTGCAAGCTCACCATTCTTTTCACTCATGATGATCATGATATTGCT 2460
2221 GAAGCACCACCACTGAACATTTGGGGCGGTAAAGCAGCAAGAGCTAAAGCTGGAGCAATTT 2280
2461 GAAGCACCACCACTGAACATTTGGGGCGGTAAAGCAGCAAGAGCTAAAGCTGGAGCAATTT 2520
2281 AAGGCCACTACAGAAATTTCTCAGCAGCAACATGCTGCTCAATTTAGACCTGGTATC 2340
2521 AAGGCCACTACAGAAATTTCTCAGCAGCAACATGCTGCTCAATTTAGACCTGGTATC 2580
2341 CAAGCAGTTGACTTGAATTTCTACTCATCTCATGATCTTCCAGCAATTTAGCATCTCGC 2400
2581 CAAGCAGTTGACTTGAATTTCTACTCATCTCATGATCTTCCAGCAATTTAGCATCTCGC 2640
2401 CAAGCAGTTGACTTGAATTTCTACTCATCTCATGATCTTCCAGCAGGGTTAACCGATCTGCA 2460
2641 CAAGCAGTTGACTTGAATTTCTACTCATCTCATGATCTTCCAGCAGGGTTAACCGATCTGCA 2700
2461 AGAAATTTTCCAGCAGCAATAGCAACCATGCAAGCGATGATGATGATGATGATGATGATGAT 2520
2701 AGAAATTTTCCAGCAGCAATAGCAACCATGCAAGCGATGATGATGATGATGATGATGATGAT 2760
2521 GCTGGACAGTAGTCTTCTGATCTTAAAGAAATCCATGCTGCGAGCGCATTTCTGAGATG 2580
2761 GCTGGACAGTAGTCTTCTGATCTTAAAGAAATCCATGCTGCGAGCGCATTTCTGAGATG 2820
2581 ATGATTCATCAATAGCAGCTTCCCACTATGGAATCTTCTAGCAGCAACCGATG 2640
2821 ATGATTCATCAATAGCAGCTTCCCACTATGGAATCTTCTAGCAGCAACCGATG 2880
2641 GAGTCTCAACTTCAATTTCTCAGTATGATCAATATCAATGATCAATGATCAATGATCAATGAT 2700
2881 GAGTCTCAACTTCAATTTCTCAGTATGATCAATATCAATGATCAATGATCAATGATCAATGAT 2940
2701 TCATATGCACTTAACCTCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2941 TCATATGCACTTAACCTCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
2761 CAGCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
3001 CAGCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
2821 TTGCTGCAAGGAAATTTGCAAACTGCTGCGAGCACTGCTGCGAGCACTGCTGCGAGCACTGCT 2880
3061 TTGCTGCAAGGAAATTTGCAAACTGCTGCGAGCACTGCTGCGAGCACTGCTGCGAGCACTGCT 3120
2881 TTAGGAGTGTCAACAGGATTAACATCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
3121 TTAGGAGTGTCAACAGGATTAACATCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
2941 CTGAATCTGGAATGTTTCAAGCAAAATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
3181 CTGAATCTGGAATGTTTCAAGCAAAATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3001 AGTCAGATTTTATCAGCGAGGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3241 AGTCAGATTTTATCAGCGAGGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
3061 AAAATGGTTTCACTTCTGATCGGTTTGTGAGACAGGATATCTGTATTAATCACTAACAAGAAC 3120
3301 AAAATGGTTTCACTTCTGATCGGTTTGTGAGACAGGATATCTGTATTAATCACTAACAAGAAC 3360
3121 CAGCTGATTTTATCAGCGAGGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3171
3361 CAGCTGATTTTATCAGCGAGGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3411

RESULT 3

AK103364
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PubMed

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

AK103364 3626 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J033126N23, full
insert sequence.

AK103364.1 GI:32988573
FULL_CDNA; CAP trapper.

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Tsunoda, Y.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hahidume, W., Hayatsu, N., Imetani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

12869764

2 (bases 1 to 3626)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirao, T.,
Hori, P., Hotta, I., Iida, Y., Iida, Y., Ikeda, R., Imamura, K.,
Imetani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagata, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Nomura, K., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602 Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

<p>Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.</p> <p>Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Tagawa, T., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Koijima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Rurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tanaka, T., Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.</p> <p>Location/Qualifiers</p>		<p>1. 3626</p> <p>/organism="Oryza sativa (japonica cultivar-group)"</p> <p>/mol_type="rRNA"</p> <p>/cultivar="Nipponbare"</p> <p>/db_xref="taxon:39947"</p> <p>/clone="J033126N23"</p>
<p>FEATURES</p> <p>source</p>		<p>Query Match 93.9%; Score 2978.4; DB 8; Length 3626;</p> <p>Best Local similarity 100.0%; Pred. No. 0;</p> <p>Matches 2979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>
Qy	1	ATGGAGATGTTGCACGTAGATCAGAGGAGCTCGTGTGTTGGGACGAACTGTATGCTT 60
Db	647	ATGGAGATGTTGCACGTAGATCAGAGGAGCTCGTGTGTTGGGACGAACTGTATGCTT 706
Qy	61	GCTGTGTGGAACTGTGTCTGTAGCGCGAGTGTGTGAGCTGACACGACGCGCTCGTCAG 120
Db	707	GCTGTGTGTGGAACTGTGTCTGTAGCGCGAGTGTGTGAGCTGACACGACGCGCTCGTCAG 766
Qy	121	GATCGACGCGCTGAAGCTGGTGTAGACAAACCGGCACACCAATGCGAGCAATTCCTCC 180
Db	767	GATCGACGCGCTGAAGCTGGTGTAGACAAACCGGCACACCAATGCGAGCAATTCCTCC 826
Qy	181	ATAAGAGGATATGTGTCTCTTCACAGAGAGATCCAAAATTCGTCTCTATCTCGG 240
Db	827	ATAAGAGGATATGTGTCTCTTCACAGAGAGATCCAAAATTCGTCTCTATCTCGG 886
Qy	241	ATTTTCCTATGACCAAAAAATGTGATGAACACAAAGCTAGTTCGAAGCCATTTCCTGTA 300
Db	887	ATTTTCCTATGACCAAAAAATGTGATGAACACAAAGCTAGTTCGAAGCCATTTCCTGTA 946
Qy	301	GCAAGTTTCGACGATGGGATGTCTGAGTGTCTGAGATGTTGGATTTGAAACTTCAGATAT 360
Db	947	GCAAGTTTCGACGATGGGATGTCTGAGTGTCTGAGATGTTGGATTTGAAACTTCAGATAT 1006
Qy	361	GGAAACAGCACCAAGAACTCTTCCCGCAAAAGCAGAAATGGCAACAGTGTGTTGCTCCATC 420
Db	1007	GGAAACAGCACCAAGAACTCTTCCCGCAAAAGCAGAAATGGCAACAGTGTGTTGCTCCATC 1066
Qy	421	ACATTTGTTGCGAGCACTTTTGTGCTGCTAGTGTGTTGGTTCACAAAAGTGTCTCTTACG 480
Db	1067	ACATTTGTTGCGAGCACTTTTGTGCTGCTAGTGTGTTGGTTCACAAAAGTGTCTCTTACG 1126
Qy	481	ACACAATCATCTCAAGGGAAGAAATGCTGATAGATCAACTCTTCCAAAAGAGTGTGCAGAA 540
Db	1127	ACACAATCATCTCAAGGGAAGAAATGCTGATAGATCAACTCTTCCAAAAGAGTGTGCAGAA 1186
Qy	541	GGCAATGACTCCAAATGCAATGCGCTTCTGGCAGAGTGGAGTCTGTGAGCGCAATACT 600
Db	1187	GGCAATGACTCCAAATGCAATGCGCTTCTGGCAGAGTGGAGTCTGTGAGCGCAATACT 1246
Qy	601	GATTTCACCAATGAAAGATTTCCAAAGGGCCACGCCAAAATATGATGTGCGAGCAAAATGTC 660

Db 5747 CTGCTAGTGTGGTTCCAAAAAGTGTCCTCTAGCACCAATCATCTCAAGGAGAGATG 5806
Qy 506 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAGCAATGATCCAAATGCAATGCGC 565
Db 5807 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAGCAATGATCCAAATGCAATGCGC 5866
Qy 566 CTCTGCGCAAGATGAGCTGCTGAGGCCAATGCTGATGCTCACCAGTGA 614
Db 5867 CTCTGCGCAAGATGAGCTGCTGAGGCCAATGCTGATGCTCACCAGTGA 5926
Qy 615 ----- 614
Db 5927 ATGTAGAGCCTTCAAAATTCCTAAGTAGGATTTTATTTAAGGTATAGATAAACTAATGT 5986
Qy 615 -----AGATTTGCAAGGCCAGGCCAAAATTTATGATGTGGCAGCAATGT 659
Db 5987 TTGTGTGATTTCTCAGATTTGCAAGGCCAGGCCAAAATTTATGATGTGGCAGCAATGT 6046
Qy 660 CTCTGAGGACAACTTCTGCTGATGTTGGGGCTTTTACCTGAAAGTTCCTCCAGATTAATG 719
Db 6047 CTCTGAGGACAACTTCTGCTGATGTTGGGGCTTTTACCTGAAAGTTCCTCCAGATTAATG 6106
Qy 720 GCACATAGAGTAATAGGTGCAGATCAAACTTCATCCACTCCCAAACTTTCTGAGTGGT 779
Db 6107 GCACATAGAGTAATAGGTGCAGATCAAACTTCATCCACTCCCAAACTTTCTGAGTGGT 6166
Qy 780 CCTCAAAGAAATGAAGATGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 839
Db 6167 CCTCAAAGAAATGAAGATGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 6226
Qy 840 CAATTTGACCAAGATCCTAAACCAATGCTGGAAGAAAGCTGATCAGTTCCTGAGCA 899
Db 6227 CAATTTGACCAAGATCCTAAACCAATGCTGGAAGAAAGCTGATCAGTTCCTGAGCA 6286
Qy 900 GTGCAATTTGACCAAGATCCTAAACCAATGCTGGAAGAAAGCTGATCAGTTCCTGCA 959
Db 6287 GTGCAATTTGACCAAGATCCTAAACCAATGCTGGAAGAAAGCTGATCAGTTCCTGCA 6346
Qy 960 TGAGCCATGTGAAGAGTGTCTCAAAGAAAGCTCCAAATCTAAGAGGAAGCGGATAA 1019
Db 6347 TGAGCCATGTGAAGAGTGTCTCAAAGAAAGCTCCAAATCTAAGAGGAAGCGGATAA 6406
Qy 1020 GAAGTTGATGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1079
Db 6407 GAAGTTGATGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6466
Qy 1080 TGCAAAAGCTTTGTCGAGAAAGCCAAAAGAGGTGCGGCTTCTATCAGAAATTAATATGC 1139
Db 6467 TGCAAAAGCTTTGTCGAGAAAGCCAAAAGAGGTGCGGCTTCTATCAGAAATTAATATGC 6526
Qy 1140 TAACCAAGTTGAGATTTCTAGAGTGAAGAGTTCATGCTGAAATGCGCGCTGATCCCTG 1199
Db 6527 TAACCAAGTTGAGATTTCTAGAGTGAAGAGTTCATGCTGAAATGCGCGCTGATCCCTG 6586
Qy 1200 TGAGGATGATAGAGTACCATCCCGTCCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 1259
Db 6587 TGAGGATGATAGAGTACCATCCCGTCCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 6646
Qy 1260 CAACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGAACAGCAAAACGCAAAATA 1319
Db 6647 CAACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGAACAGCAAAACGCAAAATA 6706
Qy 1320 CTCTGATGTTGATAGATGATGATCATCACTATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1379
Db 6707 CTCTGATGTTGATAGATGATGATCATCACTATGAGTGAAGTGAAGTGAAGTGAAGTGA 6766
Qy 1380 AACTGGAAGTGTGATCAGCAGTGTGCTCATCCAGCTGGGAAATTTGAGCAACAAAAGT 1439
Db 6767 AACTGGAAGTGTGATCAGCAGTGTGCTCATCCAGCTGGGAAATTTGAGCAACAAAAGT 6826
Qy 1440 GACACCACTGCGAGTACTGAGATGATGAGATGATGAGATGATGAGATGATGAGATGATGAG 1499

Db 6827 GACACCACTGCGAGTACTCAGCATGATGATGAGAAATGATACTGAAATGGTCTTTGACAC 6886
Qy 1500 AAATATGATATAGACAGATGCTCTGTCAGCATGTATCAGAAATCTCCACACAGAGTGCTC 1559
Db 6887 AAATATGATATAGACAGATGCTCTGTCAGCATGTATCAGAAATCTCCACACAGAGTGCTC 6946
Qy 1560 ATCAAAGGGGAAAAACAGCGGGTTTGAAGAGGGGAAAAACACATTCAGCTGCTAGTACCAA 1619
Db 6947 ATCAAAGGGGAAAAACAGCGGGTTTGAAGAGGGGAAAAACACATTCAGCTGCTAGTACCAA 7006
Qy 1620 ATATGGTGTGTAAGCAGCAGCAAAATGGTCAGAAATACATACATCTACTCAGCGGAGAGATCA 1679
Db 7007 ATATGGTGTGTAAGCAGCAGCAAAATGGTCAGAAATACATACATCTACTCAGCGGAGAGATCA 7066
Qy 1680 ATGCCAGATGGAACCCGAAACTCTGTTCTGAGTCACTCGGCAA 1723
Db 7067 ATGCCAGATGGAACCCGAAACTCTGTTCTGAGTCACTCGGCAAGGTACGAATTTGTG 7126
Qy 1724 ----- 1723
Db 7127 AATCATGAGGAATTTTGTCTTTTAAATTTGACTGAATCAACATTTTATCTGTATGAGGAA 7186
Qy 1724 ----- 1723
Db 7187 TAATATGCTGATTAACCAATGTTTAAAGAAATATGATACATGTTTATTTATATGCTTTCC 7246
Qy 1724 ----- 1723
Db 7247 ACTGTTCTTTTACTTATGTTTGTGATCTCTTTTGTGTGTGCTGATGCTGATGCT 7306
Qy 1724 ----- 1723
Db 7307 GT 7366
Qy 1724 -----AGTTT 1729
Db 7367 AGACTCATATTATAGTGAATTAAGTACATGACATTTTCTCTCATTTCTCATCTCAGGTTT 7426
Qy 1730 CTCGAGCTGAGCATGATATCCAAATTTGTGTGACTTTCATGTCAGCAGAGTCTACCCAAAG 1789
Db 7427 CTCGAGCTGAGCATGATATCCAAATTTGTGTGACTTTCATGTCAGCAGAGTCTACCCAAAG 7486
Qy 1790 AGAAAAAGAAACAAAACTTGAAGTGAATCTGCTGAAAAACAGACCATGATGATGATGATCC 1849
Db 7487 AGAAAAAGAAACAAAACTTGAAGTGAATCTGCTGAAAAACAGACCATGATGATGATGATCC 7546
Qy 1850 CCATGATATGTTGAACTGCTAGCTAATAAAACAGCATGAGAGGAGCTTTGACTGAGTGA 1909
Db 7547 CCATGATATGTTGAACTGCTAGCTAATAAAACAGCATGAGAGGAGCTTTGACTGAGTGA 7606
Qy 1910 CTGATTTGTTCTGACATCAACCGTATTTCATCCAAGACACTGCTGATGATGATTTGTGTA 1969
Db 7607 CTGATTTGTTCTGACATCAACCGTATTTCATCCAAGACACTGCTGATGATGATTTGTGTA 7666
Qy 1970 TAGTAGTCTCCAAAGGATGTTTCAAGTATTGATCAAGTGTGTTGATCACTAATTTCCCAAC 2029
Db 7667 TAGTAGTCTCCAAAGGATGTTTCAAGTATTGATCAAGTGTGTTGATCACTAATTTCCCAAC 7726
Qy 2030 AGAGTCTCTGTCATCCCAAGTATACAGAGGATTAACAGGTCATTTGGGCAATTCACCA 2089
Db 7727 AGAGTCTCTGTCATCCCAAGTATACAGAGGATTAACAGGTCATTTGGGCAATTCACCA 7786
Qy 2090 CACAAGAGTCTCCACATCTCAGAACTTTCACTCTCAGGAAACAGCAGCAGCAGCAGCAGCAG 2149
Db 7787 CACAAGAGTCTCCACATCTCAGAACTTTCACTCTCAGGAAACAGCAGCAGCAGCAGCAGCAG 7846
Qy 2150 GAATGGAAGAAATGTTCTATTTGCTGAGCTCAGCTCAGCTATTTTTCATCATCATGATGATC 2209
Db 7847 GAATGGAAGAAATGTTCTATTTGCTGAGCTCAGCTCAGCTATTTTTCATCATCATGATGATC 7906
Qy 2210 AGTATATTCTCTGAAGCAACCACTGAACTTTGGGGCCGCTTAAGGACGCAAAAGAGCTTAACT 2269
Db 7907 AGTATATTCTCTGAAGCAACCACTGAACTTTGGGGCCGCTTAAGGACGCAAAAGAGCTTAACT 7966

CDS
 join(10735...10850,10933...10987,11596...11653,12285...12379,
 12471...12574,12958...13086,13493...15302)
 /note="unnamed protein product; Similar to Arabidopsis
 thaliana chromosome 4, BAC clone F4D11; putative
 myb-protein. (AL022537)"
 /codon_start=1
 /protein_id="BAA94769.1"
 /db_xref="GI:7630236"
 /translation="MTSDNGKADPKDGEPPSPAFQGEISNEPKRRPLNGRTTGP
 TRSTKGNWTFEEDAIISRAVOTYNGKMKIAECFFDRTDVOCLRWKQVKNFELVK
 PGEKDEEDLIVQVNLKGLPKQWSTIAQALPGRIGKQCRERYWHLNPGKINAEWQ
 BEIILIAHMYNGKMAELTKFAPGRDINSIKHNSVKKVKNYSGLHTTVSCCL
 PLNYSANCOONSAMTQSEDSQFAVREVSQCSOSLAKVSCSVDHDTVPLG
 CDLVANFEDKNAHDSSQSMGPOACTSABAVASALPAVACHVSSNLDPPQHLQED
 FQGLNLDITDENPTFSPADNQTVCISIEHNSERSLEPTDVAWEVPLSMSSDSGAQ
 KLFMSADPNSCLSELQWQDLSLQGLSGPDVAEDSISNHSQSDVYSSEADTH
 FLAPYPTQNSSSVMGLADDSPQSPVPSLCSNMTDDAPFNPRGRKEMPLSQ
 AEVYTSQSSSGSVMFANPCNDNRHPSSTWESIPECCDQVNAEPASLEKEP
 SLTQSVTAPDQKAGALFYEPFSPSLDPFVSCDLVQDFEFGFPIQRLMHSITM
 NVCTPMLWSPTHESTGLVLSAASFKTPTSLIKRHRDLISLPIKRIEKKYGT
 EKDGVSDTSSTGIQTSICINATKDDALIITVLRIERSASSKSLKLVPSDENKILG
 YTTBOTKDGQSGNDEHMDROTNGERSSATNATNDDLSGLNLSSTSP"
 complement(join(18608...18823,19905...20031,20072...20220,
 20774...21113,21407...21738,21975...22508))
 /note="unnamed protein product; ESTs
 AU082761(S5084),D42006(S5084) correspond to a region of
 the predicted gene.
 Similar to Caenorhabditis elegans chromosome 3, cosmid
 C17G10; alcohol dehydrogenase/ribitol dehydrogenase.
 (U28739)"
 /codon_start=1
 /protein_id="BAA94770.1"
 /db_xref="GI:7630237"
 /translation="MAAAGAGGGGGGAYDPSVYDPSVKTFAHMYRHYRDKNYE
 IHOMYEGFORLSLPRDPAABAAVSPYCDGSHVFLLLYRELRYHAYARASSS
 SSSPAITAGCARLSWANYCDLFSVLHGVNMOLPNQMLMDWDEYVQFCQVRA
 KLNKSDDELLQPKPANNVGLNLYQALVEKSMIAHILEREKEGLEQFTADGY
 DYQGSNVLMLQVSNIGLRLHCLLDYRTGLKCLAPDIDNQOQVTVIGSHISA
 LYHGPNLMRYAALREFNKILLYLYKYQYHQSQYQDILKKNQMYAFVAVC
 LSLCPNLNDENYVLTQLEKYNKMTQKQRFDEETAAAYDELFTACPKFTPPSPFA
 LDQPLNYNDVQLQKFLYEVKQQLSLGYSYLYKDYSTITIAKLAQYMEVEDEAT
 LRAMKNSQVLYSLMTYKHKHADVNNKIVSSADFDFYIKECSSIRTKPLSKRN
 LAHIIHQVYKVRKSLSPSFSFYFFHIGNSYENGYEYRKNRKNESPAQIR"
 join(23045...23145,24592...24670,25118...25216,25302...25415,
 25608...25750,26364...26445,27073...27158,27527...27760,
 29332...29917)
 /note="unnamed protein product; Similar to Arabidopsis
 thaliana chromosome 2 BAC T19C21; unknown protein.
 (AC004683)"
 /codon_start=1
 /protein_id="BAA94771.1"
 /db_xref="GI:7630238"
 /translation="MAGSMVDRTSDMLIGPDWAKWMEICDINRDPQSGQDVYKALK
 KRIGHKPKQILLATILETAIKNCGDIEMHVAERDVLHEWVKLVKKSQDQVKEKV
 LTMIDTQEARGPARYQYVYAAHDLVAGAAFPKRSRPAFLPENGSPAGRNWRS
 PQDPERASGNDPFLMSGEIQNARGINDVIAEMNALDPGNREGLRQVIVELWD
 QCRTYQGVILLVNATDEELMSGLANDLQRLVLAQDAANGAVNREKVKLSLQ
 ALVDTSDMSQDQSKQALVDIEDPTTQETNKEPSQASQFPQGLALPAPFVNGS
 ATPAPKSDPLGLDLSMDNPSSTENSLAVPTDVPVDSPTSQNALVIDIISQNNIT
 NNTAKPADPFCVNSSTSLPGSQPYNAPTQPLQAOQPOQVQVPLPNGGAVNPTGYPT
 SSGWNGQIANNATPPAQTVNYGIALVLYIYSIQSPSYTIPIIGCQNHACFLCSV"
 30166...30675
 /note="unnamed protein product; ESTs
 AU031545(S61825),AU031546(S61825) correspond to a region
 of the predicted gene.
 hypothetical protein"
 /codon_start=1
 /protein_id="BAA94772.1"
 /db_xref="GI:7630239"
 /translation="MYNNQOPGAMQSPQAVTOMQSGFANQSGMPPHSMQMOPEGM
 OPSFWRGQOPMYAOOPMMPGQFRAAMPQPMYQPMQSOYRLVQOQRAQYVNSQGR
 PITYAGKNDLSQRYGLSMQDSSYMGNNISSPYSTTFSSSSSQGQSPKPKLFGDLS
 IATKQKQNA"

CDS

complement(join(34629...34795,36831...36867,38803...38895,
 41780...41830))
 /note="unnamed protein product; hypothetical protein"
 /codon_start=1
 /protein_id="BAA94773.1"
 /db_xref="GI:7630240"
 /translation="MTVTRCTTQRLLLEGRIQCDIRGKILSGKPPFSQWRKNTGI
 EIVGLLDACVIVITVEKPENNRIKKHGLRGLSLRCLDRLGPRLVWDSPTKVTIPIY
 VQFPRGHTLPHRG"
 join(46367...46423,47537...47769,48031...48153,50082...50299,
 51762...51864,52091...52204,52433...52532,52784...53004,
 54084...54514,54601...55709,56020...57468)
 /note="unnamed protein product; hypothetical protein"
 /codon_start=1
 /protein_id="BAA94774.1"
 /db_xref="GI:7630241"

CDS

WICAPVAAAIICAPASVPDQVPSCLRAVPPPNRQTPSPVAPSPDIDRHIPL
 ICGFMEDLLFMTHGVKQVADDEEVGRSGSARTELKIRGLVGLMLVPCGPPPIHVAGDG
 FMCVRLHVMGNGRLLALLTCSLTKANCLAMKLFPMKEKOTFELGNIIEAASG
 ERSTARVDRAPPLRRGCRSKPPSPCLSGLELLASMTKALIPHVAEEKKIPSF
 AGILAAFWNLFLAILIPARISGALIEHLEI EVDIVCACLVDLSLVVIMDEIVAV
 DOEGARVYGNCLMARGCTGAVPVELELATPRQDAAGVNDPEPAQCEHSIRGY
 VALQKDPKPSLSRIFHDQKKDBHKASSPFSVAKFARMDCSKDLKLTSDNGT
 APTLPKQNGTSDGCSITFYRSTFVASVQSVQSPSTOSSQGRNADSTLTPKSVQE
 GNSKMAPSGKGAABANTDPMKDLQCPAQNYDVAANVSDNTVDVMSGLPEVQI
 TMHLEVADQPPSTPKLSEVLRNEDNGKTEETLVAQCNLTDPMFMSGRDQ
 VAEQCNLTQPKVPQSGOKROIENPECEEVLRSSSKSKRTDKLMLKQOQSKKRTA
 QADVDAKLCKRRKAKVRLLESEINAMOVEDSRSDVRENAADPCEDDORSTIPVME
 VSMIDIPVSNHTVGEEDGLKSSKNKTKRYSDVDVDSGLSLMNLMLGKKRGTGSHVHTVAH
 PAGNLNNKVTPTASTOHDNDENTENGLDTPNMHTDVCQHVSEISQVSSKGTAGL
 SKGTHSAASTKGGESTRNGONIHYLSADQCOMETENSVLSSAKVSPAEDIIQIM
 SDHQEQLPKKKKKQKLEVTREKQTMIDIPMDIVELLAKNQHRLQMTETDCSDINR
 IQKTEADDDPVAACQSDYASSVPDNTNOSQSLASQSTQKELQGLHALITQESPH
 POFQSTQEOQTHLREEMWTIAASSPLFSEHDDQYIAEAPTEHGRKDAKLTWQF
 KATTPRQAATCGAEPGIAQVADLTSFVMSSSNYSASQPVIAPLDRAEAVQV
 HANFPQTIATMEASLCKDRNAGVVLVYPKESMPATHLLRMDDPSTLASFPNYGTS
 RNOMESOLNSOYAHKQYKGTSTQSVYSLNGLKIPLTFEDLSHOLHDLRPLRPHR
 VGTLSLLOKEIANWSENGTQSGYKLGSTGTSHTOMKREHFEALNSGMSAKNA
 IQLGSSSSSADFASNSIAQSWTRGKNVHFLDRFVQDICIITNKNPADFTIEND
 NBYMDYK"
 complement(join(58634...58667,58796...58902,59018...59083,
 59368...59432,59585...59648,60333...60442,60620...60707,
 60835...60945))
 /note="unnamed protein product; ESTs
 AU082698(R0665),D23963(R0665) correspond to a region of
 the predicted gene.
 Similar to Arabidopsis thaliana rac GTP binding protein
 Arac7. (AF079484)"

CDS

Query Match 81.1%; Score 2573.2; DB 8; Length 150594;
 Best Local Similarity 87.8%; Pred. No. 0;
 Matches 3003; Conservative 0; Mismatches 23; Indels 396; Gaps 2;
 QY 146 ACGAACGGCAGACACCAATGCGAGCATTTCTCCATAGAGGGGTATGTTCTTCTTCTTC 205
 DB 5404 ATGCATCTTTATAATGATAATCTATTTTCTTGTCACGAGGGTATGTTCTTCTTCTTC 54103
 QY 206 AGAAGAGGATCCAAAATTCGTCTCTATCTCGGATTTTCCATGACCAAGAAAATGTG 265
 DB 54104 AGAAGAGGATCCAAAATTCGTCTCTATCTCGGATTTTCCATGACCAAGAAAATGTG 54163
 QY 266 ATGACACCAAGCTAGTTCAGCCCAATTTCTCTAGCAAGTTTCCGACGATGSGATTGCT 325
 DB 54164 ATGACACCAAGCTAGTTCAGCCCAATTTCTCTAGCAAGTTTTCGACGATGSGATTGCT 54223
 QY 326 CGAAGTCTTGGATAAGTTGAAAACCTTCAGATAATGGAACAGACCAAGAACTCTTCCCG 385
 DB 54224 CGAAGTCTTGGATAAGTTGAAAACCTTCAGATAATGGAACAGACCAAGAACTCTTCCCG 54283
 QY 386 CAAGCAGATGCGACAAAGTGTGTTGTCCTCATCTTCATTTGTCGAGCACTTTTGTGC 445
 DB 54284 CAAGCAGATGCGACAAAGTGTGTTGTCCTCATCTTCATTTGTCGAGCACTTTTGTGC 54343

446 CTGCTGCTGTTGGTTCCTCCAAAAGTGTCTCTCTAGCACACATCATCTCAAGGGGAAGATG 505
Db |||||
54344 CTGCTAGTGTGGTTCCTCCAAAAGTGTCTCTCTAGCACACATCATCTCAAGGGGAAGATG 54403
Qy |||||
506 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAGCAATGATCTCCAAATGCAATGGCC 565
Db |||||
54404 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAGCAATGATCTCCAAATGCAATGGCC 54463
Qy |||||
566 CTTCTGGCAGAAATGGAGCTCTCAGGCCAATATCTGATTCACCAATGAA----- 614
Db |||||
54464 CTTCTGGCAGAAATGGAGCTCTCAGGCCAATATCTGATTCACCAATGAAAGGTATGGTAG 54523
Qy |||||
615 ----- 614
Db |||||
54524 ATGTAGAGCCTTTCAAATTCCTAAGTAGAGTATTTATTTAAGGTATAGAATAAACTAATGT 54583
Qy |||||
615 ----- AGATTTGCAAGGCCAGCCCAAAATTTATGATGTGGCAGCAATGT 659
Db |||||
54584 TTGTGTGATTTTCTCAGATTTGCAAGGCCAGCCCAAAATTTATGATGTGGCAGCAATGT 54643
Qy |||||
660 CTCTGAGCACACACTCTCTGTGATGTGGGCTTTTACCTGAAGTTCCTCCAGATTACATG 719
Db |||||
54644 CTCTGAGCACACACTCTCTGTGATGTGGGCTTTTACCTGAAGTTCCTCCAGATTACATG 54703
Qy |||||
720 GCACATAGAAATGATGTGAGATCAACCTCCATCCATCCATCCAACTTCTGAAAGTGGT 779
Db |||||
54704 GCACATAGAAATGATGTGAGATCAACCTCCATCCATCCAACTTCTGAAAGTGGT 54763
Qy |||||
780 CCTCAGAAAGATGAAGATGAATGGAATACTGAAGAGACTCTTTGTTGCTGAGCAGTG 839
Db |||||
54764 CCTCAGAAAGATGAAGATGAATGGAATACTGAAGAGACTCTTTGTTGCTGAGCAGTG 54823
Qy |||||
840 CAATTTGACCAAGATCCTTAACCCATGTCTGGAAGAGAAAGTGAATCAGTGTCTGAGCA 899
Db |||||
54824 CAATTTGACCAAGATCCTTAACCCATGTCTGGAAGAGAAAGTGAATCAGTGTCTGAGCA 54883
Qy |||||
900 GTGCAATTTGACCAAGATCCTTAACCCATGTCTGGAAGAGAAAGTGAATCAGTGTCTGAGCA 959
Db |||||
54884 GTGCAATTTGACCAAGATCCTTAACCCATGTCTGGAAGAGAAAGTGAATCAGTGTCTGAGCA 54943
Qy |||||
960 TGAGCCATGTGAGAGAGTGTCTCTCAAAAGAGCTTCCAAATCTAAGAGAGAGACGATAA 1019
Db |||||
54944 TGAGCCATGTGAGAGAGTGTCTCTCAAAAGAGCTTCCAAATCTAAGAGAGAGACGATAA 55003
Qy |||||
1020 GAAGTTGATGAAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1079
Db |||||
55004 GAAGTTGATGAAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 55063
Qy |||||
1080 TGCAAGCTTTCTCGAGAGAGCCAAAGAGTGGGCTTCTATCAGAAATTTATAAATGC 1139
Db |||||
55064 TGCAAGCTTTCTCGAGAGAGCCAAAGAGTGGGCTTCTATCAGAAATTTATAAATGC 55123
Qy |||||
1140 TAACCAAGTTGAGGATTTAGAAAGTGAAGAGTTCATCGTGAATTTGCCGTGATCCCTG 1199
Db |||||
55124 TAACCAAGTTGAGGATTTAGAAAGTGAAGAGTTCATCGTGAATTTGCCGTGATCCCTG 55183
Qy |||||
1200 TGAGATGATGAAGTACCATCCCGTCCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1259
Db |||||
55184 TGAGATGATGAAGTACCATCCCGTCCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 55243
Qy |||||
1260 CAACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1319
Db |||||
55244 CAACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 55303
Qy |||||
1320 CTCGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
Db |||||
55304 CTCGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 55363
Qy |||||
1380 AACTGGAAGTGTGCATCACACAGTGTCTCATCCAGCTGGGAATTTGAGCAACAAAAAGT 1439
Db |||||
55364 AACTGGAAGTGTGCATCACACAGTGTCTCATCCAGCTGGGAATTTGAGCAACAAAAAGT 55423

1440 GACACCCACTGCGAGTACTCAGCATGATGATGAGATGATGATGATGATGATGATGATGATGAT 1499
Db |||||
55424 GACACCCACTGCGAGTACTCAGCATGATGATGAGATGATGATGATGATGATGATGATGATGAT 55483
Qy |||||
1500 AATATGCTAATGACAGATGCTCTGTCAGCATGATCAGAAATCTCCACACAGAGGTGCTC 1559
Db |||||
55484 AATATGCTAATGACAGATGCTCTGTCAGCATGATCAGAAATCTCCACACAGAGGTGCTC 55543
Qy |||||
1560 ATCAAAGGGGAAAACAGCGGTTTGAATAGGGGAAAACACATTCAGCTGCTAGTACCAA 1619
Db |||||
55544 ATCAAAGGGGAAAACAGCGGTTTGAATAGGGGAAAACACATTCAGCTGCTAGTACCAA 55603
Qy |||||
1620 ATATGCTGTGAAAGCAGCAGCAATGCTCAGACATACATGATCTCAGCGCAGAGATCA 1679
Db |||||
55604 ATATGCTGTGAAAGCAGCAGCAATGCTCAGACATACATGATCTCAGCGCAGAGATCA 55663
Qy |||||
1680 ATGCGAGATGAAAACCCGAAACCTCTGTTCTGAGTCACTCGGCAA----- 1723
Db |||||
55664 ATGCGAGATGAAAACCCGAAACCTCTGTTCTGAGTCACTCGGCAAAGGTACGAAATTTGTG 55723
Qy |||||
1724 ----- 1723
Db |||||
55724 AATCATGAGGAATTTTGTCTTTTAAATGACTGAATCAACATTTATCTGTATGAAAGAA 55783
Qy |||||
1724 ----- 1723
Db |||||
55784 TAATATGCTGCATACAAATGTTAAGAAATATGCATACAAATGTTATTATATGCTTTCC 55843
Qy |||||
1724 ----- 1723
Db |||||
55844 ACTGTTCTCTTACTTATCTTTTGAATCTCTTTTGTGTGCGTGATGATGATGATGATGAT 55903
Qy |||||
1724 ----- 1723
Db |||||
55904 GT 55963
Qy |||||
1724 ----- AGTTTT 1729
Db |||||
55964 AGACTCATATTATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 56023
Qy |||||
1730 CTCAGCTGAGCATGATATCCAAATTTATGCTGACCTTCATGACGAGTCTACCCAAAGA 1789
Db |||||
56024 CTCAGCTGAGCATGATATCCAAATTTATGCTGACCTTCATGACGAGTCTACCCAAAGA 56083
Qy |||||
1790 AGAAAAAGAGCAAAAACTTGAAGTGTCTGTGAAAAACAGACCATGATGATGATGATGATGAT 1849
Db |||||
56084 AGAAAAAGAGCAAAAACTTGAAGTGTCTGTGAAAAACAGACCATGATGATGATGATGATGAT 56143
Qy |||||
1850 CCATGATATTTGTTGAAGTGTCTGTGAAAAACAGACCATGATGATGATGATGATGATGATGAT 1909
Db |||||
56144 CCATGATATTTGTTGAAGTGTCTGTGAAAAACAGACCATGATGATGATGATGATGATGATGAT 56203
Qy |||||
1910 CTGATTTCTGACATCAACCGTATTCATCCAAAGACCACTGCTGATGATGATGATGATGATGAT 1969
Db |||||
56204 CTGATTTCTGACATCAACCGTATTCATCCAAAGACCACTGCTGATGATGATGATGATGATGAT 56263
Qy |||||
1970 TAGTAGCTGCCAAGATGTTTCAAGTATGATCAAGTGTGTTTTCACATTAATTTCCCAAC 2029
Db |||||
56264 TAGTAGCTGCCAAGATGTTTCAAGTATGATCAAGTGTGTTTTCACATTAATTTCCCAAC 56323
Qy |||||
2030 AGAAGTCTCTGGCATCCCAAGTATCAGAGAGAGTTCAGGGTCAATTTGGCAATTTGACCA 2089
Db |||||
56324 AGAAGTCTCTGGCATCCCAAGTATCAGAGAGAGTTCAGGGTCAATTTGGCAATTTGACCA 56383
Qy |||||
2090 CACAGAGTCTCCCATCTCTCAGAACTTTTCAGTCTTACTCAGAGAACAGCAGACATTTGTC 2149
Db |||||
56384 CACAGAGTCTCCCATCTCTCAGAACTTTTCAGTCTTACTCAGAGAACAGCAGACATTTGTC 56443
Qy |||||
2150 GGATGGAAGATGCTGCTATTTCTGCAAGCTCAGCACTATTTTTCATCATGATGATGATC 2209
Db |||||
56444 GGATGGAAGATGCTGCTATTTCTGCAAGCTCAGCACTATTTTTCATCATGATGATGATC 56503
Qy |||||
2210 AGTATATTGCTGAAGCAACCAACTGAACTGAACTTGGGGCCGTTAAGGACGCAAAAGAGCTAA 2269

VERSION	166494.1	GI:2724471
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 7218)	
AUTHORS	Dorner, P., Scheiflinger, F. and Falkner, F. Gunter.	
TITLE	Recombinant fowlpox virus	
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;	
FEATURES	Location/Qualifiers	
source	1..7218	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	2.0%; Score 62.8; DB 6; Length 7218;	
Best Local Similarity	7.2%; Pred. No. 1.9e-05;	
Matches	31; Conservative 227; Mismatches 174; Indels 0; Gaps 0;	
Qy	942 ATGTGACGACAGTCTGCAATGAGCCATGTGCAAGAAGTTGTTCTCAAAGAAGCTTCCAAATC 1001	
Db	1497 ATGTAGCGATCACTGTATTAACCTATCTATGCAAGTAGTTAAAGAGATAGAGAATTTGG 1438	
Qy	1002 TAAGAGGAAGACGATAGGAAGTTGATGAAGACGACGACGACGACGACGACGACGACGACGACG 1061	
Db	1437 TACRR 1378	
Qy	1062 CCAGGCTGATGTTTCAGATGCAAGCTTTGTGCGAAGAACGACCAAAAGGTCGCGTTCT 1121	
Db	1377 RRR 1318	
Qy	1122 ATCAGAAATTATAATGCTTAACGAGTTGAGGATTTAGAACTGACGAACTTCATCGTGA 1181	
Db	1317 RRR 1358	
Qy	1182 AAATGCGCTGATCCCTGTGAGATGATAGATGATCATCCCGTCCCGATGGAAGTAAG 1241	
Db	1257 RRR 1198	
Qy	1242 CATGGATATTCCTGTAGCAACCATCAGTGGGAGAGATGGGTTAAATCAAGTAAGAA 1301	
Db	1197 RRR 1138	
Qy	1302 CAAGACAAAACGCAATACTCTGATGTTGATGATGATGATCATCACTATGAACTGGCT 1361	
Db	1137 RRR 1078	
Qy	1362 GAATGGAAAAA 1373	
Db	1077 RRRRRRRRRRA 1066	
RESULT 7		
AX655393/c	2000 bp DNA linear	PAT 22-MAR-2003
LOCUS	Sequence 5263 from Patent WO03000896.	
DEFINITION	AX655393	
ACCESSION	AX655393.1	
VERSION	AX655393.1	GI:29158207
KEYWORDS		
SOURCE	Oryza sativa	
ORGANISM	Oryza sativa	
REFERENCE		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
JOURNAL	Ehrhartoideae; Oryzaceae; Oryza.	
FEATURES		
source	1. 2000	
	/organism="Oryza sativa"	

RESULT	9
LOCUS	BX897685/c
DEFINITION	Danio rerio clone CH211-234H8, WORKING DRAFT SEQUENCE, 11 unordered pieces
ACCESSION	BX897685
VERSION	BX897685.1 GI:40067210
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

	AX655393	2000 bp	DNA	linear	PAT 22-MAR-2003
LOCUS	Sequence 5263 from Patent WO03000898.				
ACCESSION	AX655393				
VERSION	. AX655393.1 GI:29158207				
KEYWORDS	.				
SOURCE	Oryza sativa				
ORGANISM	Oryza sativa				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Erihartoidae; Oryzaceae; Oryza.				

```
misc feature
fragment_chain:1
48796..64511
/note="assembly fragment:00672
fragment_chain:1"
```

```

ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 178469)
Direct Submission
Submitted (10-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 7, 2003 this sequence version replaced gi:38141601.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZC72X22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 178208 bases at least Q40
Consensus quality: 178375 bases at least Q30
Consensus quality: 178425 bases at least Q20
Insert size: 178469; sum-of-contigs
Insert size: 166477; 18.8% error; agarose-fp
Quality coverage: 8.49x in Q20 bases; sum-of-contigs Quality
coverage: 9.18x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*   1 178469: contig of 178469 bp in length.
    Location/Qualifiers
        1..178469
         /organism="Danio rerio"
         /mol_type="Genomic DNA"
         /db_xref="taxon:7955"
         /clone="CH211-72K32"
         /clone_lib="CHORI-211"
        misc_feature          1..178469
                               /note="assembly_fragment:00026
                                clone_end:SP6
                                vector_side:left"
ORIGIN
Query Match           1.6%; Score 49.8; DB 2; Length 178469;
Best Local Similarity 44.8%; Pred.No. 0.093;
Matches 192; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

Qy 1104 AAAAAAGTCGGCTTCTATCAGAATTATAATGCTACACCAGGTGAGGATTTCTAGAAG 1163
Db 84103 AAATAATGATGAATAATAATAATAATGAATGTTTATAATATGATGATTAATAATAATAA 84162

Qy 1164 TGACGAAGTTTCATCGTGAAAATGCCCTCATCCCTGTGAGGATGATAGAAGTACCATCCC 1223
Db 84163 TGAATAATAATAATATCAATAATAATAATAATAATGATGATGATTAATAATAATAA 84222

Qy 1224 GGTCGCCGATGGAAGTAGGATATTCCTGTTAGCAACCATACAGTGGGAGAGATGG 1283
Db 84223 TGATAATAATGATTAATAATAATGATTAATAATAATTAATAATAATTTGTAATAATAA 84282

Qy 1284 GTTTAAATCAGTAAGAACAAGAACAAACGC AAAATACTCTGATGTTGTAGATGATGCATC 1343
Db 84283 TAATTATTAATAATGATGATTAATAATAATAATGATGATGATGATTAATAATAATAA 84342

Qy 1344 ATCACCTTATGAACCTGGCTGAATCGGAAAAAGAAAGAACTGGAGTGTGCATCACACAGT 1403

```


AUTHORS Pelan.S.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
COMMENT On Oct 31, 2003 this clone request: c1onerequest@sanger.ac.uk replaced gi:37665445.

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfiah-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

Only a small overlap as described above; less otherwise noted: all this sequence was finished as follows: unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >= 30), or annotated and repeats; all sequencing problems; such as plasmid subclone and repeats; all regions were covered by at least one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

occasion of the clone being a 1A6. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Racon repeat submitted.
discovery system (Zhairong Bao and Sean Eddy, submitted), and those beginning 'dr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
DKEY-43H1
is from a Zebrafish BAC library
VECTOR: pindigoBAC-5.

```

FEATURES
source
vector: pandas.Series
location/Qualifiers
1..124927
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43H1"
/clone_lib="DanioKey"

```

```

ORIGIN
Query Match 1.5%; Score 48.4; DB 5; Length 124927;
Best Local Similarity 43.7%; Pred. NO. 0.22;
Matches 214; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

```

[illegible]

```

Assembly program: Phrap; version 0.990329
Consensus quality: 220033 bases at least Q40
Consensus quality: 223217 bases at least Q30
Consensus quality: 234937 bases at least Q20
Estimated insert size: 228618; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

```

```

* * be preserved.
* * 1 227067: contig of 227067 bp in length
* * 227068 227167: gap of unknown length
* * 227168 228730: contig of 1563 bp in length.
      Location/Qualifiers
source          1..228730
                /organism="Rattus norvegicus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10116"
                /clone="CH230-14F23"
                /c186..221140
                /note="wgs contig"
misc_feature    222291..224502
                /note="wgs contig"
misc_feature    224553..227067
                /note="wgs contig"

```

```

/ note="wgs_contig"
Query Match      1.5%  Score 48.6;  DB 2;  Length 228730;
Best Local Similarity 41.6%  Pred. No. 0.2;
Matches 117;  Conservative 0;  Mismatches 164;  Indels 0;  Gaps 0;

772  GAACTGTGCTCAAAAGCAAAATGAAGATGAAATGAAAAAATCTGAAGAGACTCTCTTGCT 831
|||||
87894  GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 87953

832  GAGCAGTGCATTTTGACCAAGATCTCTAACCCATGCTCTGGAAGGAAGCTGATCAGTT 891
|||||

```

[illegible]

881134	AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	881174
RESULT 13		
BX510987	124927 bp	DNA linear VRT 31-OCT-2003
CCUCS		
DEFINITION	zebrafish DNA sequence from clone DKEY-43H1 in linkage group 20,	
	complete sequence.	
ACCESSION	BX510987	
VERSION	BX510987.7	GI:38141680
SYNWORDS	HTG.	
SOURCE	Danio rerio (zebrafish)	
ORGANISM	Danio rerio	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;	
	Cypriniformes; Cyprinidae; Danio.	
REFERENCE	1 (bases 1 to 124927)	

[illegible]

repeat_region 6718..6780 /rpt_type=tandem
repeat_region 6719..6790 /rpt_type=tandem
repeat_region 7543..7568 /rpt_type=tandem
repeat_region 7761..7816 /rpt_type=tandem
repeat_region 7824..7849 /rpt_type=tandem
repeat_region 7871..7930 /rpt_type=tandem
gene complement(<7953..>8438)
mRNA /locus tag="PF14_0240"
CDS /locus tag="PF14_0240"
/codon_start=1
/product="ribosomal protein L21e, putative"
/protein_id="AA336853.1"
/db_xref="GI:23497308"
/translations="MSGKRGKSGTGYKFKKHGECTANKYLEKLYGQYVDIV
CDSQTQGNPFNYHKTGKIPLITRGVGVLYNKRVRIRIQKKVCVRIEIVRSRC
NEDFLLKTKNAELIKKALIKNEHINIKRTEGPKPAKIKVPPSKLIITIEPLPYEE
Y"
repeat_region 8467..8550
repeat_region 8473..8539 /rpt_type=tandem
repeat_region 8477..8537 /rpt_type=tandem
repeat_region 8723..8844 /rpt_type=tandem
repeat_region 8723..8840 /rpt_type=tandem
repeat_region 8782..8844 /rpt_type=tandem
repeat_region 9019..9059 /rpt_type=tandem
repeat_region 9019..9058 /rpt_type=tandem
repeat_region 9133..9189 /rpt_type=tandem
repeat_region 9221..9254 /rpt_type=tandem
repeat_region 9604..9637 /rpt_type=tandem
repeat_region 9668..9723 /rpt_type=tandem
repeat_region 10036..10073 /rpt_type=tandem
repeat_region 10445..10472 /rpt_type=tandem
repeat_region 10511..10551 /rpt_type=tandem
gene complement(<10565..>11080)
mRNA /locus tag="PF14_0241"
CDS /locus tag="PF14_0241"
/codon_start=1
/product="basic transcription factor 3b, putative"
/protein_id="AA336854.1"
/db_xref="GI:23497309"
/translations="MEKISPEILAAAKLKKYKGNLRQIQGGKARRKIKYKHNKI
SNFKINILKLGKASVGVDEICVYRTGDTFLFKPKLCLASLOSNTYIVTGKNE
HKIDINKLFLGKGNKLDLWLEKIKNDPNIKMLKSGSDTFFKEDEQAEANDVPD
LVNFEEVSXE"
repeat_region 11429..11458 /rpt_type=tandem
repeat_region 11514..11584 /rpt_type=tandem

repeat_region /rpt_type=tandem
11514..11579
Query Match 1.5%: Score 48.4; DB 3; Length 250029;
Best Local Similarity 27.3%: Pred. No. 0.23;
Matches 94; Conservative 87; Mismatches 163; Indels 0; Gaps 0;
QY 1170 AGTTCATCGTGAATAATGCCGCTGATCCCTGTGAGAGATGATGAAGTACATCCCGTCC 1229
DB 129594 AGAGAAATCGATAAATTAATAAATAATTTGAAAGACAGACAGAGAGATTTAG 129595
QY 1230 GATGAAGTAAAGCATGGATATTCCTGTAGCAACCATACAGTGGGAGAGATGGGTAAA 1289
DB 129534 GTTAATAAGAGATGKRAAWTAGYNGRAAMCTTTKKKGAGAGATCMAANTYTKRWWNM 129475
QY 1290 ATCAAGTAAGACACAGACAAAACGCAATCTCTGATGTGTGTAGATGATGATCATCAT 1349
DB 129474 MATTWTWMAAAWAAARAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 129415
QY 1350 TATGAACGTGCTGAATGGAATAAAGAAAGAAAGTGAAGTGTGCATCACACAGTTGCTCA 1409
DB 129414 ANMKRWKRAARRARRARRAARWRRWRKRRWYTTTKRWNAANAAWKAAMSGRRMMMSRA 129355
QY 1410 TCACGTGCGAATTTGACCAACAAAAGTGCACCCACTGCGAGTACTCAGCATGATCA 1469
DB 129354 WAAARRWTTWAKRWMMWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 129295
QY 1470 TCAGATGATGATCTGAAATGCTTTCGACACAAATATGCTATAGA 1513
DB 129294 TGRAAATTATATGAT 129251
RESULT 15
AC026135 141041 bp DNA linear HTG 05-APR-2000
LOCUS Homo sapiens chromosome 3 clone RP11-78022 map 3, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.
ACCESSION AC026135
VERSION AC026135.2 GI:7417828
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 141041)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, N., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hago, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McSwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Miens, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, F., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Piller, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,
Vassiliou, B., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
TITLE
JOURNAL

COMMENT

On Apr 5, 2000 this sequence version replaced gi:7264205.
All repeats were identified using RepeatMasker;
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L7353
Center clone name: 78 O 22

Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 132847 bases at least Q40
Consensus quality: 137397 bases at least Q30
Consensus quality: 139005 bases at least Q20
Insert size: 145000; agarose-fp
Insert size: 139941; sum-of-contigs
Quality coverage: 4.8 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1
* 2190: contig of 2190 bp in length
* 2191: gap of 100 bp
* 2291: contig of 3352 bp in length
* 5642: contig of 100 bp
* 5643: gap of 100 bp
* 5743: contig of 4316 bp in length
* 10058: contig of 4316 bp in length
* 10059: gap of 100 bp
* 10158: gap of 100 bp
* 10159: contig of 4661 bp in length
* 14820: gap of 100 bp
* 14920: contig of 4489 bp in length
* 19409: gap of 100 bp
* 19509: contig of 6507 bp in length
* 26115: gap of 100 bp
* 26216: contig of 13440 bp in length
* 39655: gap of 100 bp
* 39756: gap of 100 bp
* 54129: contig of 14374 bp in length
* 54230: gap of 100 bp
* 64519: contig of 10289 bp in length
* 64619: gap of 100 bp
* 79905: contig of 15286 bp in length
* 80005: gap of 100 bp
* 109845: contig of 29841 bp in length
* 109846: gap of 100 bp
* 109946: contig of 31096 bp in length.

Location/Qualifiers
1..141041
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-78022"
/clone.lib="RPC1-11 Human Male BAC"
1..2190
/note="assembly_fragment"
2291..5642
/note="assembly_fragment"
5743..10058
/note="assembly_fragment"
10159..14819
/note="assembly_fragment"
14920..19408

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

/note="assembly_fragment"
19509..26115
/note="assembly_fragment"
26216..39655
/note="assembly_fragment"
clone_end:SP6
vector_side:right
39756..54129
/note="assembly_fragment"
54230..64518
clone_end:T7
vector_side:right
64619..79904
/note="assembly_fragment"
80005..109845
/note="assembly_fragment"
109946..141041
/note="assembly_fragment"

ORIGIN

Query Match 1.5%; Score 48.2; DB 2; Length 141041;
Best Local Similarity 46.5%; Pred. No. 0.25;
Matches 155; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
QY 1178 GTGAAATGCGGTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATCGAAG 1237
Db 98342 GAGATCATGCGCATTTACACTCCAGCTGGTCAACAGAGCGAGATTCCATCTCAAAAAA 98401
QY 1238 TAACATGGATATTCCTGTTAGCAACCATACATGAGGAGAGATGGTTAAATCAAGTA 1297
Db 98402 AAAAAAGTAAATGGCATTAGGAAAAACATTGTAAAGGAGAGAGGTAATATGTTACAA 98461
QY 1298 AGAACAAAGACAAACGCAAAATCTCTGTATGTTGTAGATGATGATCATCATCTTATGA 1357
Db 98462 ATACAAATCAAAACCAACACAGTTAGACTTAAAAAAGATATTAAGAGATTGAAAC 98521
QY 1358 GGCTGAATGGAAGAAAGCAAGAAAGAACTGGAGTGTGATCATCACAGTTGCTCATCCAGCTG 1417
Db 98522 CTGATGGAGAAATCTGTAAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 98581
QY 1418 GGAAATTTGAGCAACAAAAAGTACACCCCTCGAGTACTCGAGCATGATGATGAGATG 1477
Db 98582 GACATTGAGGCAACAAAAAGCAAAAAATTTTAAAAATTACAAATTATCAAAATATTCTCG 98641
QY 1478 ATACTGAAATGGTCTTGACACAAATATGCATA 1510
Db 98642 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 98674

Search completed: May 16, 2004, 23:34:15
Job time : 12119 secs

FEATURES
source

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2004, 15:00:26 ; Search time 1184 Seconds
(without alignments)
11377.564 Million cell updates/sec

Title: US-09-721-114-1_COPY_655_3825

Perfect score: 3171
Sequence: 1 atggacatgttcagtaga.....acgagtatatggattaccgc 3171

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3171	100.0	3896	7	Abq77908 Rice OsEM
2	3171	100.0	4310	6	Abq93084 Oryza sat
3	2573.2	81.1	9455	6	Abq93085 Oryza sat
4	73.4	2.3	279	6	Ab173385 Corn tass
5	62	2.0	2000	7	Ada71938 Rice gene
6	52	1.6	2000	7	Ada71938 Rice gene
7	45.4	1.4	274	7	Abx85499 Corn ear-
8	42.8	1.3	574	3	Aac44195 Arabidops
9	42.2	1.3	110000	6	Abq67194 Listeria
10	42.2	1.3	319630	6	Abq67194 Listeria
11	41.4	1.3	246	5	Aas90679 DNA encod
12	41.4	1.3	1269	6	Abn90734 Staphyloc
13	41.4	1.3	2456	5	Aas75925 DNA encod
14	41.4	1.3	2848	5	Aah54196 S. epider
15	41	1.3	1260	7	Ac446331 Prokaryot
16	40.6	1.3	297	6	Abn90752 Staphyloc
17	40.6	1.3	5299	7	Aad49594 Human cyt
18	40.6	1.3	5322	7	Aad49592 Human cyt
19	40.6	1.3	5451	9	Adc26234 Human NOV
20	40.6	1.3	5761	6	Abz11141 Human pol
21	40.4	1.3	686	3	Aaa09318 Human can
22	40.4	1.3	110000	6	AbA03041_17
23	40	1.3	412	8	ACH35151 Human end

C	24	39.8	1.3	15122	6	ABL32775	Human imm
C	25	39.6	1.2	3257	4	AAH54703	S. epider
C	26	39.4	1.2	843	5	AAS92080	DNA encod
C	27	39.2	1.2	2522	2	AAV63196	CDNA from
C	28	39.2	1.2	2522	6	ABQ92046	Human pol
C	29	39.2	1.2	6065	6	ABL32504	Human imm
C	30	38.8	1.2	648	7	ABT42404	Toxicity
C	31	38.8	1.2	648	9	ADB59062	Toxicity
C	32	38.8	1.2	648	9	ADB53815	Primary r
C	33	38.8	1.2	3291	7	ABZ42128	Arabidops
C	34	38.8	1.2	14041	4	AAH48024	Internal
C	35	38.6	1.2	261	4	AAI27301	Probe #17
C	36	38.6	1.2	261	4	AAI22038	Probe #11
C	37	38.6	1.2	261	4	ABA67114	Human foe
C	38	38.6	1.2	261	4	ABA75585	Human foe
C	39	38.6	1.2	261	4	AAI56167	Probe #24
C	40	38.6	1.2	261	4	AAI47330	Probe #16
C	41	38.6	1.2	261	4	ABA49199	Human bre
C	42	38.6	1.2	261	4	ABA34204	Probe #12
C	43	38.6	1.2	261	4	ABA40197	Probe #18
C	44	38.6	1.2	261	4	AAK41287	Human bon
C	45	38.6	1.2	261	4	AAK50216	Human bon

ALIGNMENTS

RESULT 1
ABQ77908
ID ABQ77908 standard; cDNA; 3896 BP.
AC ABQ77908;
XX
XX
DT 24-JAN-2003 (first entry)
DE Rice OsEMF1-encoding cDNA, SEQ ID NO:1.
XX
XX
KW Rice; OsEMF1; EMF; embryonic flower; plant; floral repressor;
KW reproductive development; flower development; transgenic plant;
KW antisense suppression; transgenic; reciprocal negative interaction;
KW flower meristem identity gene; flowering time; shoot development;
KW seed yield; agriculture; gene; ss.
XX
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
CDS 241..3414
FT /*tag= a
FT /product= "OsEMF1 protein"

XX WO200280659-A1.
XX 17-OCT-2002.
XX 05-APR-2002; 2002WO-US012675.
XX 06-APR-2001; 2001US-00828068.
XX (BEGC) UNIV CALIFORNIA.
XX Moon Y, Chen L, Sung ZR;
XX WPI; 2003-046831/04.
XX P-PSDB; ABB99878.
XX New isolated OsEMF1 nucleic acid molecule and encoded polypeptide, useful
XX for plant genetic engineering, in particular controlling reproductive
XX development in rice.
XX Claim 3; Fig 1; 47pp; English.
XX The invention relates to a rice EMF (embryonic flower) gene designated
XX OsEMF1 (cDNA given in ABQ77908) and its encoded protein (ABB99878). The

Db 1981 CATGATATCAAAATTAATGCTGACCTTCAATGACGAGTCTACCCAGAGAGAAAAAGAG 2040
Qy 1801 CAAAACTTGAAGTCACTGCTGAAAAAAGACAGACATGATAGATGACATCCCATGATAT 1860
Db 2041 CAAAACTTGAAGTCACTGCTGAAAAAAGACAGACATGATAGATGACATCCCATGATAT 2100
Qy 1861 GTTGAATCTGCTAGCTATAAAACAGACATGAGAGGAGCTTATGACTGAGCTGATGTTCT 1920
Db 2101 GTTGAATCTGCTAGCTATAAAACAGACATGAGAGGAGCTTATGACTGAGCTGATGTTCT 2160
Qy 1921 GACATCAACCGTATTAATCAATCCCAAGACAACTGCTGATGATGATGTTGTAATAGTACTGCT 1980
Db 2161 GACATCAACCGTATTAATCAATCCCAAGACAACTGCTGATGATGATGTTGTAATAGTACTGCT 2220
Qy 1981 AAGSAGTGTTCAGATTAATCAATCAAGTGTGTTTGAACATAATCCCAACAGAGCTCTTG 2040
Db 2221 AAGSAGTGTTCAGATTAATCAATCAAGTGTGTTTGAACATAATCCCAACAGAGCTCTTG 2280
Qy 2041 GCATCCCAAGTACACAGAGAGGTACAGGGTCAATTTGGCAATGACCAACAGAGTCT 2100
Db 2281 GCATCCCAAGTACACAGAGAGGTACAGGGTCAATTTGGCAATGACCAACAGAGTCT 2340
Qy 2101 CCACATCTCTCAGAACTTTTCAGTCTACTCAGAGAACAGCAGACACATTTGGGATGGAAGAA 2160
Db 2341 CCACATCTCTCAGAACTTTTCAGTCTACTCAGAGAACAGCAGACACATTTGGGATGGAAGAA 2400
Qy 2161 ATGGTCACTATTGCTGCAAGCTCAACATATTTTCAATCATCATGATGATCATGATATGCT 2220
Db 2401 ATGGTCACTATTGCTGCAAGCTCAACATATTTTCAATCATCATGATGATCATGATATGCT 2460
Qy 2221 GAAGCACAACCTGAACTTTGGGCGCTTAAGGACGCAAGAGCTAACGTTGGGAGCAATTT 2280
Db 2461 GAAGCACAACCTGAACTTTGGGCGCTTAAGGACGCAAGAGCTAACGTTGGGAGCAATTT 2520
Qy 2281 AAGGCCACTACAGAAATTTCTCAGCAGCAAAATGCTGCTCAATTTAGACCTGCTATC 2340
Db 2521 AAGGCCACTACAGAAATTTCTCAGCAGCAAAATGCTGCTCAATTTAGACCTGCTATC 2580
Qy 2341 CAAGCAGTGTGCTGCTGCTTCTACTCATGTCATGGATCTTCAGCAATTTATGCTCATCTGC 2400
Db 2581 CAAGCAGTGTGCTGCTGCTTCTACTCATGTCATGGATCTTCAGCAATTTATGCTCATCTGC 2640
Qy 2401 CAACCAAGTAAATTCGCGCACTGCAACCGCTATGCTGAAAGAGCGGTTAACAGGTCATGCA 2460
Db 2641 CAACCAAGTAAATTCGCGCACTGCAACCGCTATGCTGAAAGAGCGGTTAACAGGTCATGCA 2700
Qy 2461 AGAAATTTCCAGGCAATATGCAACCATGAGAGGAGTAAGTATGATGCTGCGGAAT 2520
Db 2701 AGAAATTTCCAGGCAATATGCAACCATGAGAGGAGTAAGTATGATGCTGCGGAAT 2760
Qy 2521 GCTGGACAAGTATGCTGCTATCTTAAGAAATCCATGCTGCGACGCTCTTCTGGAATG 2580
Db 2761 GCTGGACAAGTATGCTGCTATCTTAAGAAATCCATGCTGCGACGCTCTTCTGGAATG 2820
Qy 2581 ATGGATCCATCAACATGAGCAAGCTTCCCACTATGGAATCTTAGCAGGAAACAGATG 2640
Db 2821 ATGGATCCATCAACATGAGCAAGCTTCCCACTATGGAATCTTAGCAGGAAACAGATG 2880
Qy 2641 GAGTCTCAACTTCATTAATTTCTCAGTATGACATAATCAGTACAAAGGATCAACCGACACA 2700
Db 2881 GAGTCTCAACTTCATTAATTTCTCAGTATGACATAATCAGTACAAAGGATCAACCGACACA 2940
Qy 2701 TCATATGGCAGTAACCTGATGAAAGATTCATGACATTCGAGACTTATCAGCGCAT 2760
Db 2941 TCATATGGCAGTAACCTGATGAAAGATTCATGACATTCGAGACTTATCAGCGCAT 3000
Qy 2761 CAGTGCATGATCTGCACAGACCTTTACGCCACATCTTAGAGTGTGCTGCTGCTCC 2820
Db 3001 CAGTGCATGATCTGCACAGACCTTTACGCCACATCTTAGAGTGTGCTGCTGCTCC 3060
Qy 2821 TTGCTGCAGAGAAATTCGAACTGGTCCGGAACCTGCGCAACAATCTGCTTATAG 2880
Db 3061 TTGCTGCAGAGAAATTCGAACTGGTCCGGAACCTGCGCAACAATCTGCTTATAG 3120

Qy 2881 TTAGAGTGTCAACAGGAATAACATCCATCAGATGAACAGAAAGGAACATTTTGAAGCC 2940
Db 3121 TTAGAGTGTCAACAGGAATAACATCCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Qy 2941 CTGAATTTCTGAAATGTTTTCAGCAAAATGGAATGCAATTCGAGTTGGTTCCTGTTAGCTCC 3000
Db 3181 CTGAATTTCTGAAATGTTTTCAGCAAAATGGAATGCAATTCGAGTTGGTTCCTGTTAGCTCC 3240
Qy 3001 AGTCGAGATTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACAGAGGCAAGGCT 3060
Db 3241 AGTCGAGATTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACAGAGGCAAGGCT 3300
Qy 3061 AAAATGCTTCACTCCCTGGATCGGTTTGTGAGCAGGATATCTGTATTAACATAAGAAC 3120
Db 3301 AAAATGCTTCACTCCCTGGATCGGTTTGTGAGCAGGATATCTGTATTAACATAAGAAC 3360
Qy 3121 CCAGTCAATTTTACTACAATCAGTAAACGATAACGAGTATATGATATCCGC 3171
Db 3361 CCAGTCAATTTTACTACAATCAGTAAACGATAACGAGTATATGATATCCGC 3411

RESULT 2

ABA93084
ID ABA93084 standard; cDNA; 4310 BP.
XX ABA93084;
AC ABA93084;
XX 11-APR-2002 (first entry)
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein encoding cDNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 655..3828
FT /*tag= a
FT /product= "brassinosteroid response related protein"
XX
XX JP2001327287-A.
XX 27-NOV-2001.
XX
XX 19-MAY-2000; 2000JP-00149106.
XX
XX 19-MAY-2000; 2000JP-00149106.
XX
XX (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
XX (SEIB-) SEIBUTSUKI TOKUTEI SANGYO GIJUTSU.
XX
XX WPI; 2002-135739/18.
XX P-PSDB; ABB05428.
XX
XX New gene involved in brassinosteroid responses useful for controlling the effects such as growth promotion, increased yield, quality improvement, ripeness promotion, stress relaxation and chemical resistance.
XX
XX Disclosure; Page 6-10; 19pp; Japanese.
XX
XX The present invention describes a polynucleotide encoding a plant gene which can control the signal transfer system of brassinosteroid hormone. The polynucleotide can be used for controlling the effects such as growth promotion, increased yield, quality improvement, ripeness promotion, stress relaxation and chemical resistance. The present sequence encodes a brassinosteroid response related protein isolated from rice (Oryza sativa), from the present invention
XX
XX Sequence 4310 BP; 1237 A; 986 C; 1022 G; 1065 T; 0 U; 0 Other;

Query Match 100.0%; Score 3171; DB 6; Length 4310;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGATTTGCAATGATATCAGAGGAGCTCTGTTGTTGGACAACTGTATGCTT 60
DB 655 ATGGAGATTTGCAATGATATCAGAGGAGCTCTGTTGTTGGACAACTGTATGCTT 714

QY 61 GCTCTGCTGGAACTGGTCTGCTAGCGCCAGTGTGGAGCTGACAGCGACGCTCGTCAG 120
DB 715 GCTCTGCTGGAACTGGTCTGCTAGCGCCAGTGTGGAGCTGACAGCGACGCTCGTCAG 774

QY 121 GATGAGCGCTGTAAGCTGTGTAGACGAAACCGGACAAACACCAATGCGAGATTTCTCC 180
DB 775 GATGAGCGCTGTAAGCTGTGTAGACGAAACCGGACAAACACCAATGCGAGATTTCTCC 834

QY 181 ATGAGAGGATGTTGCTCTCTTCCAGAGAGGATCCAAATTCCTGCTCTCTAFTCTCGG 240
DB 835 ATGAGAGGATGTTGCTCTCTTCCAGAGAGGATCCAAATTCCTGCTCTCTAFTCTCGG 894

QY 241 ATTTTCCATGACCAAGAAATGTGATGACCAAAAGCTAGTTCAAGCCCATTTTCTGTA 300
DB 895 ATTTTCCATGACCAAGAAATGTGATGACCAAAAGCTAGTTCAAGCCCATTTTCTGTA 954

QY 301 GCAAGATTTGCAAGATGGTGTCTGAGTGTCTGATATGATGATGATGATGATGATGAT 360
DB 955 GCAAGATTTGCAAGATGGTGTCTGAGTGTCTGATATGATGATGATGATGATGATGAT 1014

QY 361 GGAACAGCACCAAGAACTCTTCCGCAAGCAGAGATGGCACAAGTGTGTTGCTCCATC 420
DB 1015 GGAACAGCACCAAGAACTCTTCCGCAAGCAGAGATGGCACAAGTGTGTTGCTCCATC 1074

QY 421 ACATTTGTCGAGACATTTTGTGCTGCTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 480
DB 1075 ACATTTGTCGAGACATTTTGTGCTGCTAGTGTGTTGTTGTTGTTGTTGTTGTTG 1134

QY 481 ACACATCATCTCAAGGGAAGATGCTGATGATCAATCTTTCCAAAGAGTGTGCAAGAA 540
DB 1135 ACACATCATCTCAAGGGAAGATGCTGATGATCAATCTTTCCAAAGAGTGTGCAAGAA 1194

QY 541 GGAATGACTCCAAATGCATGCGCTTCTGCAAGATGAGCTGCTGAGCGCAATACT 600
DB 1195 GGAATGACTCCAAATGCATGCGCTTCTGCAAGATGAGCTGCTGAGCGCAATACT 1254

QY 601 GATTCACCAATGAAAGATTTGCAAGGGCCAGCGCCCAAAATATGATGTGCGACAAATGTC 660
DB 1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCGCCCAAAATATGATGTGCGACAAATGTC 1314

QY 661 TCTGAGGACAACTTCTGTTGATGTTGGGCTTTACCTGAAAGTTCCTGAGATTAATGTC 720
DB 1315 TCTGAGGACAACTTCTGTTGATGTTGGGCTTTACCTGAAAGTTCCTGAGATTAATGTC 1374

QY 721 CACATGAGATTAATGTCAGATCAACCTCCATCCATCCAAACTTTCTGAAAGTGTTC 780
DB 1375 CACATGAGATTAATGTCAGATCAACCTCCATCCATCCAAACTTTCTGAAAGTGTTC 1434

QY 781 CTCAAAGGAATCAAGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 840
DB 1435 CTCAAAGGAATCAAGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1494

QY 841 AATTTGACCAAGATCTTAACCAATGCTGGAAGGACGATGATGATGATGATGATGATG 900
DB 1495 AATTTGACCAAGATCTTAACCAATGCTGGAAGGACGATGATGATGATGATGATGATG 1554

QY 901 TGCATTTTGACCAAGATCCGAAACCAAGTGTCTGGGCAAGAAATGTGAGCAGATCTGCAAT 960
DB 1555 TGCATTTTGACCAAGATCCGAAACCAAGTGTCTGGGCAAGAAATGTGAGCAGATCTGCAAT 1614

QY 961 GAGCCATGTGAGAGATGTTCTCAAGAGAGTCCAAATCTTAAGAGGAGAGCGGATAG 1020
DB 1615 GAGCCATGTGAGAGATGTTCTCAAGAGAGTCCAAATCTTAAGAGGAGAGCGGATAG 1674

QY 1021 AAGTTGATGAAGAGCAGAGCAGCAGCAGAGAAACGACCTGCCAGGCTGATGTTTCAGAT 1080
DB 1675 AAGTTGATGAAGAGCAGAGCAGCAGCAGAGAAACGACCTGCCAGGCTGATGTTTCAGAT 1734

QY 1081 GCAAGCTTTGTCGAGAGAGCCAAAGAGCTGCGGCTTCTATCAGAAATTTATAAATGCT 1140
DB 1735 GCAAGCTTTGTCGAGAGAGCCAAAGAGCTGCGGCTTCTATCAGAAATTTATAAATGCT 1794

QY 1141 AACAGGTTGAGATTTAGAGTGACGAAAGTTCACTGCTGAGAAATGCCGCTGATCCCTGT 1200
DB 1795 AACAGGTTGAGATTTAGAGTGACGAAAGTTCACTGCTGAGAAATGCCGCTGATCCCTGT 1854

QY 1201 GAGGATGATAGAGTACCATCCCGTCCCGATGGAAGTAAAGCATGATATTTCTGTTAGC 1260
DB 1855 GAGGATGATAGAGTACCATCCCGTCCCGATGGAAGTAAAGCATGATATTTCTGTTAGC 1914

QY 1261 AACCATACAGTGGGAGAGATGGGTTAAATCAAGTAAAGAACAGACAAACCGCAATATC 1320
DB 1915 AACCATACAGTGGGAGAGATGGGTTAAATCAAGTAAAGAACAGACAAACCGCAATATC 1974

QY 1321 TCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1975 TCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2034

QY 1381 ACTGGAAAGTGTGATCACAAGTGTCTGTCATCCAGCTGGGAATTTGAGCAACAAAAGATG 1440
DB 2035 ACTGGAAAGTGTGATCACAAGTGTCTGTCATCCAGCTGGGAATTTGAGCAACAAAAGATG 2094

QY 1441 ACACCCACTGCGAGTACTCAGCATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 2095 ACACCCACTGCGAGTACTCAGCATGATGATGATGATGATGATGATGATGATGATGATG 2154

QY 1501 AATATGATGAAGACAGATGTCGTGTCAGCATGTCATCAGAAATCTCCACACAGAGGTGCTCA 1560
DB 2155 AATATGATGAAGACAGATGTCGTGTCAGCATGTCATCAGAAATCTCCACACAGAGGTGCTCA 2214

QY 1561 TCAGAGGGGAAAACAGCGGGTTTGATGTAAGGGGAAAACACATTCAGCTGCTGATACCAA 1620
DB 2215 TCAGAGGGGAAAACAGCGGGTTTGATGTAAGGGGAAAACACATTCAGCTGCTGATACCAA 2274

QY 1621 TATGCTGCTGAAGACACAGAAATGGTCAGAAATATGATGATGATGATGATGATGATGATG 1680
DB 2275 TATGCTGCTGAAGACACAGAAATGGTCAGAAATATGATGATGATGATGATGATGATGATG 2334

QY 1681 TGCCAGATGGAACCGGAAAACCTGTTCTGATGTCATGCGGAAAAGTTTCTCCAGCTGAG 1740
DB 2335 TGCCAGATGGAACCGGAAAACCTGTTCTGATGTCATGCGGAAAAGTTTCTCCAGCTGAG 2394

QY 1741 CATGATATCCAAATTTATGTCGTGACCTTCATGAGCAGAGTCTACCCAGAGAGAAAAGAG 1800
DB 2395 CATGATATCCAAATTTATGTCGTGACCTTCATGAGCAGAGTCTACCCAGAGAGAAAAGAG 2454

QY 1801 CAAAAAATTTGAGTGTCTGTAAGAAAACAGACCAATGATGATGATGATGATGATGATGATG 1860
DB 2455 CAAAAAATTTGAGTGTCTGTAAGAAAACAGACCAATGATGATGATGATGATGATGATGATG 2514

QY 1861 GTTGACCTGCTAGCTTAAACCCAGCATGAGAGGAGCTTATGATGATGATGATGATGATGATG 1920
DB 2515 GTTGACCTGCTAGCTTAAACCCAGCATGAGAGGAGCTTATGATGATGATGATGATGATGATG 2574

QY 1921 GACATCAACCGTATTTCAATCCAAGACAACTGCTGATGATGATGATGATGATGATGATGATG 1980
DB 2575 GACATCAACCGTATTTCAATCCAAGACAACTGCTGATGATGATGATGATGATGATGATGATG 2634

QY 1981 AAGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
DB 2635 AAGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2694

QY 2041 GCATCCCAAGATGACAGAGAGGATTAACAGGCTCATTTGGCAATTTGACACCAAGAGTCT 2100
DB 2695 GCATCCCAAGATGACAGAGAGGATTAACAGGCTCATTTGGCAATTTGACACCAAGAGTCT 2754

QY 2101 CCACATCTCTCAAGAACTTTTCACTGCTGATGATGATGATGATGATGATGATGATGATG 2160

2755 CCACATCTCTCAGAACTTTTCAGTCTACTCAGGAACAGCAGACACATTTTGGGATGGAAGA 2814
2161 ATGGTCACTATTGCTCGAGCTCACCACATTTTTCATCATCATGATCATCATGATATTTGCT 2220
2815 ATGGTCACTATTGCTCGAGCTCACCACATTTTTCATCATCATGATCATCATGATATTTGCT 2874
2221 GAAGCACCAACTGAACATTTGGGCGCTTAAGAGCGCAAGAGCTTAACGTTGGAGCAATTT 2280
2875 GAAGCACCAACTGAACATTTGGGCGCTTAAGAGCGCAAGAGCTTAACGTTGGAGCAATTT 2934
2281 AAGGCCACTACAGAAATTTCTCCAGCAGCAACATGTTGGTCTCAATTTAGACTGCTGATC 2340
2935 AAGGCCACTACAGAAATTTCTCCAGCAGCAACATGTTGGTCTCAATTTAGACTGCTGATC 2994
2341 CAAGCAGTTGACTTGTCTTCTACTCTATGTCATGGGATCTTCCAGCAATTTATGTCATCTGC 2400
2995 CAAGCAGTTGACTTGTCTTCTACTCTATGTCATGGGATCTTCCAGCAATTTATGTCATCTGC 3054
2401 CAACAGTATTTGGCGACTTGGCCCTATGCTGAAGAGCGGTAAACAGGTCCATGCA 2460
3055 CAACAGTATTTGGCGACTTGGCCCTATGCTGAAGAGCGGTAAACAGGTCCATGCA 3114
2461 AGAAATTTTCCAAAGCAATAGCAACCATGGAAGCGAGTAAGTTATGTCATCGGAGAAAT 2520
3115 AGAAATTTTCCAAAGCAATAGCAACCATGGAAGCGAGTAAGTTATGTCATCGGAGAAAT 3174
2521 GCTGGAAGAGTATGTTGATCTTAAAGAAATCATGCTGCGACCGCATCTTCTGAGAAATG 2580
3175 GCTGGAAGAGTATGTTGATCTTAAAGAAATCATGCTGCGACCGCATCTTCTGAGAAATG 3234
2581 ATGATCTCATACATTAAGCACTTCCCAACTATGGAACCTTCTAGCAGGAACCAAGATG 2640
3235 ATGATCTCATACATTAAGCACTTCCCAACTATGGAACCTTCTAGCAGGAACCAAGATG 3294
2641 GAGTCTCAACTTCATATTTCTAGTATGCAATATCATGTAACAAAGATCAACAGCA 2700
3295 GAGTCTCAACTTCATATTTCTAGTATGCAATATCATGTAACAAAGATCAACAGCA 3354
2701 TCATATGGCAGTAACTGTAAGTAAGTATCCATTCACATTTGCAAGCTTATCAGGCAAT 2760
3355 TCATATGGCAGTAACTGTAAGTAAGTATCCATTTGACATTTGCAAGCTTATCAGGCAAT 3414
2761 CAGCTGATGATCTGCACAGACCTTTTACGCCACATCTCTAGAGTTGGTGTGCTTGGCTCC 2820
3415 CAGCTGATGATCTGCACAGACCTTTTACGCCACATCTCTAGAGTTGGTGTGCTTGGCTCC 3474
2821 TTGCTGAGAGAAATTTGAAACTGGTGGAGAACTGTGGCAACAATCTGTTTATAG 2880
3475 TTGCTGAGAGAAATTTGAAACTGGTGGAGAACTGTGGCAACAATCTGTTTATAG 3534
2881 TTAGGAGTGTCAACAGGAATAACAATCGCATGCAAGCAAGAAATTTTGAAGCC 2940
3535 TTAGGAGTGTCAACAGGAATAACAATCGCATGCAAGCAAGAAATTTTGAAGCC 3594
2941 CTGAATTTCTGGAATGTTTTCAGCAAAATGGAATGCAATTCAGTTGGGTTCTGTTAGCTCC 3000
3595 CTGAATTTCTGGAATGTTTTCAGCAAAATGGAATGCAATTCAGTTGGGTTCTGTTAGCTCC 3054
3001 AGTGCAGATTTTATCAGCAGAGCAAGCATAGCTCAATCTTGGACCGAGCGAGGCT 3060
3655 AGTGCAGATTTTATCAGCAGAGCAAGCATAGCTCAATCTTGGACCGAGCGAGGCT 3714
3061 AAAATGGTTTATCCCTTGGATCGGTTTGTGAGCAGGATATCTGTATAACTAACAGAAC 3120
3715 AAAATGGTTTATCCCTTGGATCGGTTTGTGAGCAGGATATCTGTATAACTAACAGAAC 3774
3121 CCAGCTGATTTTACTACATTCAGTACGATAACAGGATATGATGATACGCT 3171
3775 CCAGCTGATTTTACTACATTCAGTACGATAACAGGATATGATGATACGCT 3825

ABA93085
ID ABA93085 standard; DNA; 9455 BP.
XX
AC ABA93085;
XX
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein related DNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ds.
XX
OS Oryza sativa.
XX
PN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PF 19-MAY-2000; 2000JP-00149106.
XX
PR 19-MAY-2000; 2000JP-00149106.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
XX
RA (SEIB-) SEIBUTSUXEI TOKUTEI SANGYO GIJUTSU.
XX
WP 1; 2002-135739/18.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
PS Disclosure; Page 13-16; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence
CC represents a DNA sequence related to a brassinosteroid response related
CC protein isolated from rice (Oryza sativa), from the present invention
XX
SQ Sequence 9455 BP; 2571 A; 2001 C; 2090 G; 2793 T; 0 U; 0 Other;
Query Match 81.1%; Score 2573.2; DB 6; Length 9455;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 3003; Conservative 0; Mismatches 23; Indels 396; Gaps 2;
QY 146 ACGAACCGGCGCAACACCAATCGGAGCATTTCTCCATGAAGAGGATGTTGCTCTCTTC 205
Db 5447 ATGCATCTTTTATAATGATAATCTTATTTTCTGTACAGAGGGTATGTTGCTCTCTTC 5506
QY 206 AGAACAGGATCCAAATTTCTGCTCTATCTCGATTTTCCATGACCAAGAAATGTTG 265
Db 5507 AGAAGAGATCCAAATTTCTGCTCTATCTCGATTTTCCATGACCAAGAAATGTTG 5566
QY 266 ATGAACAACAAGCTAGTTCAAGCCCATTTCTGTAGCAAAAGTTTCGAGATGGGATGCT 325
Db 5567 ATGAACAACAAGCTAGTTCAAGCCCATTTCTGTAGCAAAAGTTTCGAGATGGGATGCT 5626
QY 326 CGAGCTGCTTGGATTAAGTTGAATCTTCAGATAATGGAACAGCAACAAGAACTCTTCCCG 385
Db 5627 CGAGCTGCTTGGATTAAGTTGAATCTTCAGATAATGGAACAGCAACAAGAACTCTTCCCG 5686
QY 386 CAAGCAGATGGCAAAAGTATGTTGCTTCCATCAATTTTTCGGAGCACTTTTGTGC 445
Db 5687 CAAGCAGATGGCAAAAGTATGTTGCTTCCATCAATTTTTCGGAGCACTTTTGTGC 5746
QY 446 CTGCTAGTGTGGTTCGCCAAAGTGTCTCTTAGCACAATCTCTCAAGGAGCAATG 505
Db 5747 CTGCTAGTGTGGTTCGCCAAAGTGTCTCTTAGCACAATCTCTCAAGGAGCAATG 5806

6887	DB	AAATATGATGATGACGATGCTGTGACATGTATCAGAAATCTCCACAGAGGTGCTC	6946
1560	QY	ATCAAAAGGGGAAAAACAGCGGGTTTCAGTAAGGGGAAAAACAACTCAGTGTAGTACCAA	1619
6947	DB	ATCAAAAGGGGAAAAACAGCGGGTTTCAGTAAGGGGAAAAACAACTCAGTGTAGTACCAA	7006
1620	QY	ATATGGTGGTGAAGCACCAAGAAATGGTCCAGAACATCATGTACTCAGCGCAGAGATCA	1679
7007	DB	ATATGGTGGTGAAGCACCAAGAAATGGTCCAGAACATCATGTACTCAGCGCAGAGATCA	7066
1680	QY	ATGCCAGATGGAAATCCGAAACTCTGTGTTCTGAGTCACTCGGCAA	1723
7067	DB	ATGCCAGATGGAAATCCGAAACTCTGTGTTCTGAGTCACTCGGCAAAGGTAAGAAATTTGTGTG	7126
1724	QY	-----	1723
7127	DB	AATCATGAGGAATTTTGTCTTTTAAATTTGACTGAATCAACATTTATCTGTATGAAGGAA	7186
1724	QY	-----	1723
7187	DB	TAATATTTGGTGCATAACAATGTTAAGAAATATGCATACAATGTTTATTTATATGCTTTCC	7246
1724	QY	-----	1723
7247	DB	ACTGTTCTTTTACTTATATGTTTTCATACTCTTTTGTGTGTCGTGCATGTGTGCATGT	7306
1724	QY	-----	1723
7307	DB	GTGTGTGTGTGTGTGTGCGCGGTGTGTGTGCAGGTGCGTGGCGCAATATTTCTTTT	7366
1724	QY	-----	1723
7367	DB	AGACTCATATTATGTGATTGTAATGGACTGACATTTTCTCTCATTTCTCATCTCAGGTTT	7426
1730	QY	CTTCAGCTGAGCATGATATCCAAATATGTCTGACCTTCATGACGAGAGTCTTACCCAGA	1789
7427	DB	CTTCAGCTGAGCATGATATCCAAATATGTCTGACCTTCATGACGAGAGTCTTACCCAGA	7486
1790	QY	AGAAAAAGGAAGCAAAAACTTGAAGTGAATCTGTGAAAAAACAACCATGATGATGATCC	1849
7487	DB	AGAAAAAGGAAGCAAAAACTTGAAGTGAATCTGTGAAAAAACAACCATGATGATGATCC	7546
1850	QY	CAATGATATTGTTGAACTGTCTAGCTATAAAACCAAGCATGAGAGGAGCTTATGATGAGA	1909
7547	DB	CAATGGATATTGTTGAACTGTCTAGCTATAAAACCAAGCATGAGAGGAGCTTATGATGAGA	7606
1910	QY	CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTCATGATGATTTGTGTA	1969
7607	DB	CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTCATGATGATTTGTGTA	7666
1970	QY	TATGATGTCGCAAGGATGTTTCAGATTATGATCAAGTGTGTTTGAACACTAAATTTCCAAAC	2029
7667	DB	TATGATGTCGCAAGGATGTTTCAGATTATGATCAAGTGTGTTTGAACACTAAATTTCCAAAC	7726
2030	QY	AGAACTCTTGGCANTCCCAAGTACACAGAGGAGTTACAGGTCATTTGGCATTGACCA	2089
7727	DB	AGAACTCTTGGCANTCCCAAGTACACAGAGGAGTTACAGGTCATTTGGCATTGACCA	7786
2090	QY	CAACAGAGTCTCCCATCTCTCAGAACTTTCAGTCTTACTCAGGAAACAGCAGACACATTTGC	2149
7787	DB	CAACAGAGTCTCCCATCTCTCAGAACTTTCAGTCTTACTCAGGAAACAGCAGACACATTTGC	7846
2150	QY	GGATGGAGAAATGCTCATATTGTCGCAAGTCTCACCTATTTTTCATCATGATGATC	2209
7847	DB	GGATGGAGAAATGCTCATATTGTCGCAAGTCTCACCTATTTTTCATCATGATGATC	7906
2210	QY	AGTATATTGCTGAAGCACCACACTGAAACATTTGGGGCCGTAAGGACGCAAGAGCTAACT	2269
7907	DB	AGTATATTGCTGAAGCACCACACTGAAACATTTGGGGCCGTAAGGACGCAAGAGCTAACT	7966
2270	QY	GGAGGCAATTTAAGGCCACTACAGAAATTTCTCAGCAGACCACTGTGTGTCTCAATTTA	2329

7967	GGAGGCAATTTAAGGCCACTACAAGAAATCTCCAGCAGCAACATGTGGTGTCAATTTA	8026
2330	GACCTGGTATCCAAAGCAGTTGACTTGCATCTCTACTCATGTCAATGGGATCTTCAGCAAT	2389
8027	GACCTGGTATCCAAAGCAGTTGACTTGCATCTCTACTCATGTCAATGGGATCTTCAGCAAT	8086
2390	ATGCATCTGGCCBAACACAGTAATTTGCCCACTGACCGCTATGCTGGAAGGCGGTAAACC	2449
8087	ATGCATCTGGCCBAACACAGTAATTTGCCCACTGACCGCTATGCTGGAAGGCGGTAAACC	8146
2450	AGGTCCATGCAAGAAATTTTCCAAGCACAAATAGCAACCAATGGAAGCGAGTAAGTTATGTG	2509
8147	AGGTCCATGCAAGAAATTTTCCAAGCACAAATAGCAACCAATGGAAGCGAGTAAGTTATGTG	8206
2510	ATCGGAGAAATGCTGCAAGTAGTCTTGTAAGCTTAAAGATCCATGCGCTGGAGCGCATC	2569
8207	ATCGGAGAAATGCTGCAAGTAGTCTTGTAAGCTTAAAGATCCATGCGCTGGAGCGCATC	8266
2570	TTCTGAGATGATGGATCCATCAACATTAAGCAAGCTTCCCACACTATGCAACTCTTAGCA	2629
8267	TTCTGAGATGATGGATCCATCAACATTAAGCAAGCTTCCCACACTATGCAACTCTTAGCA	8326
2630	GGAAACAGATGGAGTCTCAACTTCATTAATTTCTCAGTAGTGCAATATCAGTACAAAGGAT	2689
8327	GGAAACAGATGGAGTCTCAACTTCATTAATTTCTCAGTAGTGCAATATCAGTACAAAGGAT	8386
2690	CAACAGCAGCATCATATGCGCTGTAACCTGAATGGAAGATTCATTTGCATTCGAGACT	2749
8387	CAACAGCAGCATCATATGCGCTGTAACCTGAATGGAAGATTCATTTGCATTCGAGACT	8446
2750	TATCAAGCGCATCAGCTGCATGATCTGCACAGACCTTTACGCCACATCTCTAGAGTTGGTG	2809
8447	TATCAAGCGCATCAGCTGCATGATCTGCACAGACCTTTACGCCACATCTCTAGAGTTGGTG	8506
2810	TGCTTGGCTCCTTGTGCGAGAGGAAATTTGCAAACTGGTGGGAGAACTCTGGCACACAAAT	2869
8507	TGCTTGGCTCCTTGTGCGAGAGGAAATTTGCAAACTGGTGGGAGAACTCTGGCACACAAAT	8566
2870	CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGGAACAGAAAGGAAC	2929
8567	CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGGAACAGAAAGGAAC	8626
2930	ATTTTGAGCCCTGGAATCTTGGATGTTTTCAGCAAAATGGAATGCAATTCAGTTGGGTT	2989
8627	ATTTTGAGCCCTGGAATCTTGGATGTTTTCAGCAAAATGGAATGCAATTCAGTTGGGTT	8686
2990	CTGTTAGCTCCAGTCGAGATTTTATTCAGCGAGGAACAGCATAGCTCAATCTTGGACCA	3049
8687	CTGTTAGCTCCAGTCGAGATTTTATTCAGCGAGGAACAGCATAGCTCAATCTTGGACCA	8746
3050	GAGCGAAGGTAAAAATGGTTCAFCCTTGGATCGGTTGTGAGACAGGATATCTGTATAA	3109
8747	GAGCGAAGGTAAAAATGGTTCAFCCTTGGATCGGTTGTGAGACAGGATATCTGTATAA	8806
3110	CTTAACAGAACCCAGCTGATTTTACTACAATCTGATTAACGATACGAGTATATGGATTACC	3169
8807	CTTAACAGAACCCAGCTGATTTTACTACAATCTGATTAACGATACGAGTATATGGATTACC	8866
3170	GC 3171	
8867	GC 8868	
RESULT 4		
ABL73385		
ABL73385 standard; cDNA; 279 BP.		
XX	ABL73385;	
XX	AC	
XX	AT	
XX	DE	
XX	14-MAY-2002 (first entry)	
XX	Corn tassal-derived polynucleotide (cdps) SEQ ID NO:2759.	

Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs; inheritance; characteristic; growth; development; disease resistance; environmental adaptability; quality; yield; molecular marker; multigene trait; plant breeding; corn tassel; gene; ss.

Zea mays.

US2001051335-A1.

13-DEC-2001.

16-APR-1999; 99US-00294093.

21-APR-1998; 98US-0082567P.

(LALG/) LALGUDI R V.

(ITOL/) ITO L Y.

(SHER/) SHERMAN B K.

Lalgudi RV, Ito LY, Sherman BK;

WPI; 2002-163647/21.

Novel purified corn tassel-derived polynucleotide useful for determining altered gene expression, to recover regulatory elements and to follow inheritance of desirable characteristics through hybrid breeding programs.

Claim 1; SEQ ID NO 2759; 201pp; English.

The present sequence describes a purified corn tassel-derived polynucleotide sequence (cdps) comprising a nucleic acid sequence selected from those given in ABL70627 to ABL76833. The cdps sequences encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I) can be used for determining altered gene expression, to recover regulatory elements and to follow inheritance of desirable characteristics through hybrid breeding programs. (I) are also useful in the evaluation, and alteration of desired characteristics associated with growth and development, disease resistance, environmental adaptability, quality and yield, and as molecular markers for studying inheritance of multigene traits in a plant breeding program. (I) can be used to produce a tassel-specific profile of gene transcription, a transcript image, to clone regulatory elements for use in transformation vectors, to express a polypeptide, to identify, isolate or extend identical or related corn tassel nucleic acid sequences from DNA libraries, in nucleic acid hybridisation or amplification technologies, as query sequences to determine homology of known sequences, as probe for use in Southern or Northern hybridisation, and to identify the presence of and/or to determine the degree of similarity between two (or more) nucleic acid sequences

Sequence 279 BP; 70 A; 74 C; 73 G; 56 T; 0 U; 6 Other;

Query Match 2.3%; Score 73.4; DB 6; Length 279;

Best Local Similarity 71.3%; Pred. No. 2.6e-10;

Matches 122; Conservative 0; Mismatches 45; Indels 4; Gaps 2;

QY 2720 ATGGAAGATTTCATTTGACATTGGAAGACTTATCAGCGCATCA---GCTGCATGATCTGC 2776

Db 105 ATGGAAGCCACCGCTGANACTGGAAGACTTCTCTCGGCGCTCATTNCACGACGACTTGC 164

QY 2777 ACAGACACTTTAGCCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCCTGCAGAGGAAA 2836

Db 165 GAGAGCCCTTTAGCNCCTCACCTCGANTCGGTGTGCTCGGTTCATTGCTGCACGACGAGA 224

QY 2837 TTGCAAACTGGTTCGGAGAACTGTGGCACACAAATCTGTGTTATAGTTAGGAG 2887

Db 225 TCGCAAACTGGTCTGGGAA-TSCGGGCCGACGCTCTGGGTANAGACTAGGTG 274

RESULT 5

ADA71938/c

ID ADA71938 standard; DNA; 2000 BP.

WP	Sequence split into 31 fragments	LOCUS ABQ69245	Accession	ABQ69245
WP	Fragment Name	Begin	End	
WP	ABQ69245_00	1	110000	
WP	ABQ69245_01	100001	210000	
WP	ABQ69245_02	200001	310000	
WP	ABQ69245_03	300001	410000	
WP	ABQ69245_04	400001	510000	
WP	ABQ69245_05	500001	610000	
WP	ABQ69245_06	600001	710000	
WP	ABQ69245_07	700001	810000	
WP	ABQ69245_08	800001	910000	
WP	ABQ69245_09	900001	1010000	
WP	ABQ69245_10	1000001	1110000	
WP	ABQ69245_11	1100001	1210000	
WP	ABQ69245_12	1200001	1310000	
WP	ABQ69245_13	1300001	1410000	
WP	ABQ69245_14	1400001	1510000	
WP	ABQ69245_15	1500001	1610000	
WP	ABQ69245_16	1600001	1710000	
WP	ABQ69245_17	1700001	1810000	
WP	ABQ69245_18	1800001	1910000	
WP	ABQ69245_19	1900001	2010000	
WP	ABQ69245_20	2000001	2110000	
WP	ABQ69245_21	2100001	2210000	
WP	ABQ69245_22	2200001	2310000	
WP	ABQ69245_23	2300001	2410000	
WP	ABQ69245_24	2400001	2510000	
WP	ABQ69245_25	2500001	2610000	
WP	ABQ69245_26	2600001	2710000	
WP	ABQ69245_27	2700001	2810000	
WP	ABQ69245_28	2800001	2910000	
WP	ABQ69245_29	2900001	3010000	
WP	ABQ69245_30	3000001	3111208	

Query Match 1.3%; Score 42.2; DB 6; Length 110000;
 Best Local Similarity 47.0%; Pred. No. 7.8; Indels 0; Gaps 0;
 Matches 131; Conservative 0; Mismatches 148; Length 110000;

WP	Sequence split into 31 fragments	LOCUS ABQ69245	Accession	ABQ69245
WP	Fragment Name	Begin	End	
WP	ABQ69245_00	1	110000	
WP	ABQ69245_01	100001	210000	
WP	ABQ69245_02	200001	310000	
WP	ABQ69245_03	300001	410000	
WP	ABQ69245_04	400001	510000	
WP	ABQ69245_05	500001	610000	
WP	ABQ69245_06	600001	710000	
WP	ABQ69245_07	700001	810000	
WP	ABQ69245_08	800001	910000	
WP	ABQ69245_09	900001	1010000	
WP	ABQ69245_10	1000001	1110000	
WP	ABQ69245_11	1100001	1210000	
WP	ABQ69245_12	1200001	1310000	
WP	ABQ69245_13	1300001	1410000	
WP	ABQ69245_14	1400001	1510000	
WP	ABQ69245_15	1500001	1610000	
WP	ABQ69245_16	1600001	1710000	
WP	ABQ69245_17	1700001	1810000	
WP	ABQ69245_18	1800001	1910000	
WP	ABQ69245_19	1900001	2010000	
WP	ABQ69245_20	2000001	2110000	
WP	ABQ69245_21	2100001	2210000	
WP	ABQ69245_22	2200001	2310000	
WP	ABQ69245_23	2300001	2410000	
WP	ABQ69245_24	2400001	2510000	
WP	ABQ69245_25	2500001	2610000	
WP	ABQ69245_26	2600001	2710000	
WP	ABQ69245_27	2700001	2810000	
WP	ABQ69245_28	2800001	2910000	
WP	ABQ69245_29	2900001	3010000	
WP	ABQ69245_30	3000001	3111208	

Query Match 1.3%; Score

XX	Homo sapiens.	
QS	WO200175067-A2.	
XX	11-OCT-2001.	
XX	30-MAR-2001; 2001WO-US008631.	
XX	31-MAR-2000; 2000US-00540217.	
XX	23-AUG-2000; 2000US-00649167.	
XX	(HYSE-) HYSEQ INC.	
XX	Dmanac RT, Liu C, Tang YT;	
XX	WPI; 2001-639362/73.	
XX	P-PSDB; ABG26492.	
XX	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity.	
XX	Claim 1; SEQ ID NO 26483; 103pp; English.	
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
CC	sequences. (I) is useful as hybridisation probes, polymerase chain	
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
CC	and in recombinant production of (II). The polynucleotides are also used	
CC	in diagnostics as expressed sequence tags for identifying expressed	
CC	genes. (I) is useful in treat disease states involving (II). (II) is	
CC	useful for generating antibodies against it, detecting or quantitating a	
CC	polypeptide in tissue, as molecular weight markers and as a food	
CC	supplement. (II) and its binding partners are useful in medical imaging	
CC	of sites expressing (II). (I) and (II) are useful for treating disorders	
CC	involving aberrant protein expression or biological activity. The	
CC	polypeptide and polynucleotide sequences have applications in	
CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	and to produce other types of data and products dependent on DNA and	
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic	
CC	coding sequences of the invention. Note: The sequence data for this	
CC	patent did not appear in the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 246 BP; 126 A; 20 C; 77 G; 23 T; 0 U; 0 Other;	
QS	Query Match 1.3%; Score 41.4; DB 5; Length 246;	
	Best Local Similarity 60.0%; Pred. NO. 0.39; Mismatches 46; Indels 0; Gaps 0;	
	Matches 69; Conservative 0;	
QY	938 AGAATGTTGAGCAGACTCTGCATTCAGCCATGTGAGAGAGTTGTTCTCTCAAAAGAGCTCCA 997	
Db	71 AGGACTGTGACAGCAAGAAGAAGAGAGGACTCGGAGAGAGAGAGAGAGAGAGAGAGA 130	
QY	998 AATCTAAGGAGGAGACCGGATAGAGAGTTGATGAGAGCAGCAGCAGCAGAGAA 1052	
Db	131 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 185	
RESULT 12		
ID	ABN90734	
ID	ABN90734 standard; DNA; 1269 BP.	
AC	ABN90734;	
AC	ABN90734;	
DT	24-JUL-2002 (first entry)	
XX	Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:197.	
XX		

KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX antibacterial; gene therapy; gene; ds.
OS Staphylococcus epidermidis.
PN US6380370-B1.
PD 30-APR-2002.
XX 13-AUG-1998; 98US-00134001.
XX 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Doucette-Stamm LA, Bush D;
XX WPI; 2002-381255/41.
PI P-PSDB; AB938189.
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
XX Disclosure; SEQ ID NO 197; 267pp; English.
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
XX Sequence 1269 BP; 478 A; 181 C; 245 G; 365 T; 0 U; 0 Other;
SQ Query Match 1.3%; Score 41.4; DB 6; Length 1269;
Best Local Similarity 49.8%; Pred. No. 1;
Matches 105; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 473 CTCTAGCACACATCTCTCAGGAGAGATCTGATAGATCACTCTTCCAGAGTG 532
Db 1058 CTATAAGTGAAGCAAGCAATCGAACGTAACAGGAGCGGTGATACGTTCAATTATAG 1117
QY 533 TCGAAGAAGCAATGACTCCAAATGCAATGCGCTTCTGCGAAGATGGAGTGTGAGS 592
Db 1118 AAGAGCATTAATTGACATCATGTATGATGATCTTCTTCTGAAATGTAAGTAAATAG 1177
QY 593 CCAATCTGATTCACCAATGAAGAATTTGCAAGGCGCCAGCCCAAAATTTATGATGTGCGAG 652
Db 1178 TCATTACTGAACAAACAAATTAATGAAGAAATTGAGCGCTGAATTATATGATGATGAAGGAA 1237
QY 653 CAAATGCTCTGAGGACCAACACTTCTGTGA 683
Db 1238 ATCTAATTATAAGAACAGACATCTGCTTA 1269
RESULT 13
AA575925
ID AA575925 standard; CDNA; 2456 BP.
XX AC AA575925;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #11729.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX

OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
PI P-PSDB; ABG11738.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 11729; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostic, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 2456 BP; 656 A; 575 C; 654 G; 571 T; 0 U; 0 Other;
Query Match 1.3%; Score 41.4; DB 5; Length 2456;
Best Local Similarity 60.0%; Pred. No. 1.5;
Matches 69; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 938 AGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTTCTCAAGAGCTCCA 997
Db 2281 AGGACTGTGAAGAGAGAAAGAGAGGACTGCGAAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 998 ATCTTAAGGAG 1052
Db 2341 AGAAG 2395
RESULT 14
AAH54196/c
ID AAH54196 standard; DNA; 2848 BP.
XX AC AAH54196;
XX 03-SEP-2001 (first entry)
XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3560.
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
XX

XX endocarditis; ds.
XX Staphylococcus epidermidis.
XX WO200134809-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US030782.
XX 09-NOV-1999; 99US-0164258P.
XX (GLAXO) GLAXO GROUP LTD.
XX Kimmerly WJ;
XX WPI; 2001-316495/33.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis.
XX Claim 8; Page 1144-1145; 2188pp; English.
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX (II), given in ARG81454 to ARG81456, from Staphylococcus epidermidis. (I)
XX and (II) can have antibacterial activity and therefore can be used in
XX vaccination. The nucleic acids (II) may be used to produce the S.
XX epidermidis polypeptides (II) via the production of vectors containing
XX them which are used to produce hosts cells which express the
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX used to vaccinate subjects and to raise antibodies against the bacteria.
XX The polypeptides may also be used to assay for other inhibitors of their
XX activity and therefore identify compounds that may be used for the
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX polynucleotide sequences from the present invention. AAH55091 to AAH55098
XX represent oligonucleotide sequences and primers which are used in the
XX exemplification of the present invention. N.B. The present invention
XX specifically claims all the polynucleotide sequences given in the
XX sequence listing of the present specification, however the sequence
XX listing only goes up to SEQ ID NO:4454 so even though sequences are given
XX in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
XX for SEQ ID NO:4455 to 4464
XX Sequence 2848 BP; 876 A; 501 C; 442 G; 1029 T; 0 U; 0 Other;
Query Match 1.3%; Score 41.4; DB 4; Length 2848;
Best Local Similarity 49.8%; Pred. No. 1.6; Indels 0; Gaps 0;
Matches 105; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 473 CTCTAGCACCAATCATCTCAAGGGAAGAATGCTGATGATCACTCTTCCAAAGAGTG 532
DB 2021 CTATAAGTGAAMAAGCAATCGAACGTAAACAGGAGCGGTGATTAAGTTTCAATTATAG 1962
QY 533 TGCAAGAGGCATGACTCCAAATGCAATGCGGCTTCTGCGAAGATGGCTGCTGAGG 592
DB 1961 AGAAGCATTAATTTGACATCATGTATGATGTACCTCTTCTGGAATGTAAGTAAGTAG 1902
QY 593 CCATATCTGATTCACCAATGAAGAATTGCAAGGCGGAGCCAGCCAAATTTATGTTGGCAG 652
DB 1901 TCATTATGACAAACAAATTAATGAAGAATTTGAGCTGATATATATGATGATGAAGAA 1842
QY 653 CAATGCTCTGAGGACACACATCTGTTGA 683
DB 1841 ATCTAATTAATGAACAACAGACATCTGCTTA 1811
RESULT 15
ACA46931
ID ACA46931 standard; DNA; 1260 BP.
XX AC
XX ACA46931;
XX

DT 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #28588.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Staphylococcus epidermidis.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029925/02.
XX P-FSDB; ASU43061.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 34801; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX product's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1260 BP; 471 A; 181 C; 244 G; 364 T; 0 U; 0 Other;
Query Match 1.3%; Score 41; DB 7; Length 1260;
Best Local Similarity 49.8%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy	473	CTCCTAGCACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTG	532
Db	1052	CTATAAGTGAAAAAGCAATCGAAGGTAAACAGGAGGCGGTGATTACGTTCAATTATAG	1111
Qy	533	TGCAAGAAGGCRATGACTCCAAATGCAATCGCGCTTCTGCAAGATGGAGCTGCTGAGG	592
Db	1112	AAGAAGCATTAATTGACATCATGTATGATGTACCTTCTTCTGAAAATGTAAGTAAAGTAG	1171
Qy	593	CCAATACTGATTCAACCAATGAAAGATTTCAGAGGCGCAGCCCAAAATTTATGATGTGGCAG	652
Db	1172	TCATTACTGACAAACAATTAATGAGAAATTGAGGCTGATTTATATGATGATGAGGGA	1231
Qy	653	CAAAATCTCTTGAGGACACACTTCTGTT	681
Db	1232	ATCTAATTAATAAGAACAGACATCTGCT	1260

Search completed: May 16, 2004, 20:12:21
Job time : 1204 secs

OM nucleic - nucleic search, using sw model

Run on: May 16, 2004, 19:21:07 ; Search time 219 Seconds
(without alignments)
8035.385 Million cell updates/sec

Title: US-09-721-114-1_COPY_655_3825

Perfect score: 3171
Sequence: 1 atggagattgttcagtaga.....acgagtatattgattaccgc 3171

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 692709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	62.8	2.0	7218	1	US-08-232-463-14
2	45.4	1.4	274	4	Sequence 14, Appli
3	41.4	1.3	1269	4	Sequence 3959, Ap
4	40.6	1.3	297	4	Sequence 197, App
5	40.4	1.3	686	4	Sequence 215, App
6	39	1.2	832	4	Sequence 9, Appli
7	38.2	1.2	335	4	Sequence 2813, Ap
8	38	1.2	19619	4	Sequence 45, Appl
9	37.4	1.2	508	4	Sequence 809, App
10	37.4	1.2	1407	4	Sequence 14, Appl
11	36.8	1.2	576	4	Sequence 535, App
12	36.8	1.2	8648	4	Sequence 2195, Ap
13	36.8	1.2	17341	4	Sequence 6914, Ap
14	36.4	1.1	2277	1	Sequence 2, Appli
15	36.4	1.1	2277	1	Sequence 2, Appli
16	36.4	1.1	2277	1	Sequence 2, Appli
17	36.4	1.1	2277	1	Sequence 2, Appli
18	36.2	1.1	1140	3	Sequence 32, Appl
19	36.2	1.1	1140	3	Sequence 32, Appl
20	36.2	1.1	1525	2	Sequence 1, Appli
21	36.2	1.1	1525	2	Sequence 1, Appli
22	36.2	1.1	1526	3	Sequence 3, Appli
23	36.2	1.1	1674	2	Sequence 1, Appli
24	36.2	1.1	1674	3	Sequence 1, Appli
25	36.2	1.1	1674	4	Sequence 1, Appli
26	36.2	1.1	1674	5	Sequence 1, Appli
27	36.2	1.1	1937	3	Sequence 7, Appli

28	36.2	1.1	1997	3	US-08-510-133A-34	Sequence 14, Appli
29	36.2	1.1	1997	4	US-09-355-700-7	Sequence 7, Appli
30	36.2	1.1	1997	4	US-08-601-132-32	Sequence 32, Appli
31	36.2	1.1	1997	4	US-08-671-573B-32	Sequence 1, Appli
32	36.2	1.1	2031	4	US-08-706-054A-1	Sequence 2, Appli
33	36.2	1.1	2031	4	US-08-706-054A-2	Sequence 1, Appli
34	36.2	1.1	2031	4	US-09-313-299-1	Sequence 2, Appli
35	36.2	1.1	2031	4	US-09-313-299-2	Sequence 1, Appli
36	35.8	1.1	289	3	US-09-007-005-17	Sequence 17, Appli
37	35.8	1.1	289	3	US-09-244-796-17	Sequence 17, Appli
38	35.8	1.1	489	4	US-09-134-001C-2451	Sequence 2451, Ap
39	35.8	1.1	2673	2	US-09-519-232-73	Sequence 73, Appli
40	35.8	1.1	11283	2	US-08-603-753D-3	Sequence 3, Appli
41	35.8	1.1	11283	3	US-09-099-753-3	Sequence 3, Appli
42	35.8	1.1	11283	3	US-08-986-106-3	Sequence 3, Appli
43	35.8	1.1	11385	2	US-08-639-501-1	Sequence 1, Appli
44	35.8	1.1	11385	2	US-09-044-946-1	Sequence 1, Appli
45	35.8	1.1	11385	3	US-09-044-908-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: PALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgt-F15
US-08-232-463-14

Query Match 2.0%; Score 62.8; DB 1; Length 7218;

	Query Match	1.4%	Score 45.4;	DB 4;	Length 274;
	Best Local Similarity	68.2%;	Pred. NO. 0.00057;		
	Matches 107;	Conservative	0;	Mismatches 46;	Indels 4; Gaps 3;
QY	2720	ATGGAAGAGATTCCATTGACATTCGAAGACTTATTCACGGCATCAGCTGCATGA--TCTGCA	2777		
DB	105	ATGGAAGCCCAACGGCTGCACATCTGGAAGACTTGTCTCGCGCTCATTTTCCACGACAGACTTGC	164		
QY	2778	CAGACCTTTTACGCCCATCTCTAGAGTTGGTGTGCTTGGCTCCCTCTGTCGAGAGAAAT	2837		
DB	165	GCAGCCCTTTACGSCCTTCACCTCGTTCGCTGTGCTCGGTTTCA-TGCTCGACGACAGAGAT	223		

1776	GAGTCTACCCAGAGAAAGAAAGACCAAAATCTGAGTGTGCTCCGTGAAAGACAGACCAT	1835
Qy		
532	ACTCCAGCCATGACAAAGAAAGACTCGAGACTTCAAGCAGTCAAAAGAGCATCGAAG	591
Db		
1836	GATAGATGACATGCCCATGGATATTTGTGAACTGCTAGCTAAAAACCCAGCATGAGAGGCCA	1895
Qy		
532	GTTCCTTAAATTCAGTCTCAGTATGAAAGACNATTGAAGAAATTCAGCAGGATGTGAT	651
Db		
1896	GCTTATGACTGAGAC	1910
Qy		
652	CGAAATGAAGAAAC	666
Db		

RESULT 6

US-09-621-976-2813/c
; Sequence 2813, Application US/09621976

Query Match

Best Local Similarity 11.2%; Pred. No. 0.13;
Matches 37; Conservative 159; Mismatches 133; Indels 3; Gaps 1;

Q-Y 1309 AAACGCAATACTCTGATGTTGTAGATGATGGATCATCACTTATGAACCTGGCTGAATGGA 1368

[illegible]

336 ARRWYYWKSTYACASRYRKYTNGWVWYMWKRMSTRWYCYMCWKCCMYRGRRCAWY TMA 277

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

1369 AAAAAGAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGC 1428
QY QY

276 RGRMWSYAGKWKSMR SAMSMCTBMYKKGSTYWTNKCTCATWCYWYWKYKRMWSKTCW 217

DD
Z76 KGRWBSIAWNNWNSRSGABNOCIFUFIHAGSUAITHTACACCHT90CINWAKKXESBEMHJGCH

1429 AACAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAGATGATGACTGAAAAT 1488
QY

[illegible]

216 SGSRGGYMTSYTSTRSYSMYWASWMYTMCWGWSTYWMWCKKWRYATTWRRAM 157

1480 CATTCTCCACCAATN - - TCCATACACACACATCTCTCAGCATGTATCAGAAATCTCC 1545

QY
I489 GGTCTTGGACACAATA---TGCAATAGGACAGAATGTCIGTCAGCATGTATCAGGAATCTCC

156 WWWWAAWTTMMNNYMWAAWCMSSRGGAAMYRRRTMMWGYYWRRKKSYYRTRCAWAYAWKTKRS 97

22

1546 ACACAGAGTGCTCATCAAGGGGAAACAGCGGGTTGAGTAAGGGGAAACACATTCA 1605

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

DB 96 Y Y W C W R W K W K R C M M M M M M M A M A Y G K T M M R A C W K T R Y R R W W A W A M R R M W W I T M M M Y Y W Y Z W R 37

1606 GGTGCTAGTACCAAAATATGGTGGTGAAAGCA 1636

[illegible]

Db 36 AMICRWWNRKWRKRSWSWMTWMAWGMTTRWAAARW 6

RESULT 7

US-09-894-844-45
Sequence 45 Application US/09894844

Sequence 45, Application US/09897844
Patent No. 6686166

; FALCILE NO: 0000100
; GENERAL INFORMATION:

APPLICANT: Behr, Marcel

APPLICANT: Small, Peter

; APPLICANT: Schoolnik, Gary

```

Query Match          1.2%; Score 38; DB 4; Length 639;
Best Local Similarity 50.5%; Pred.No.0.23;
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY      1100 AGCCAAAAGAGTGGCGCTTCTATCAGAAATTAATAATGCTAACCCAGGTTGAGGATTTCTA 1159
DB      53  AACCAAAAAGAGTGAATATTATATGCGCAGCAATTAATCAACCAATCCAGAAATTTTCAA 1112

QY      1160 GAAGTGACGAAGTTTCATCGTGAAAAATCCCGCTGATCCCTGTGAGGATGATAGAGTACCA 1219
DB      113  CCCAAGCCTATGTGAATGGTGAGTTTAAACCATCACTTCTGTGATGTTTAAAGGCGAGCT 172

QY      1220 TCCCGGTCCCATGGAAGTAAGCATGATGATATTCCTGTTAGCAACCATACATACAGTGGGAGAAG 1279
DB      173  GGGCGATTTTCATGTTTATCCACATGATTTTACTTTTGTGTTGCCCAACAGAGCTAGAGG 232

QY      1280 AT 1281
DB      ||
      233 AT 234

RESULT 9
US-09-596-002-14
; Sequence 14, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 14
; TYPE: DNA
; LENGTH: 19619
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 14
; PUBLICATION INFORMATION:
US-09-596-002-14

Query Match          1.2%; Score 38; DB 4; Length 19619;
Best Local Similarity 50.5%; Pred.No.2.8;
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY      1100 AGCCAAAAGAGTGGCGCTTCTATCAGAAATTAATAATGCTAACCCAGGTTGAGGATTTCTA 1159
DB      7061 AACCAAAAAGAGTGAATATTATGCGCAGCAATTAATCAACCAATCCAGAAATTTTCAA 7120

QY      1160 GAAGTGACGAAGTTTCATCGTGAAAAATCCCGCTGATCCCTGTGAGGATGATAGAGTACCA 1219
DB      7121 CCCAAGCCTATGTGAATGGTGAGTTTAAACCATCACTTCTGTGATGTTTAAAGGCGAGCT 7180

QY      1220 TCCCGGTCCCATGGAAGTAAGCATGATATTCCTGTTAGCAACCATACATACAGTGGGAGAAG 1279
DB      7181 GGGCGATTTTCATGTTTATCCACATGATTTTACTTTTGTGTTGCCCAACAGAGCTAGAGG 7240

QY      1280 AT 1281
DB      ||
      7241 AT 7242

RESULT 10
US-09-221-017B-535/c
; Sequence 535, Application US/09221017B
; Patent No. 6444799

```

	Query Match	1.2%;	Score 37.4;	DB 4;	Length 508;
	Best Local Similarity	47.6%;	Pred. No. 0.29;		
	Matches 110;	Conservative	0;	Mismatches 121;	Indels 0; Gaps 0;
QY	758	CTCCAAAACCTTTCTGAAGTGGTCTCTCAAAAGAAATGAAGATGAATATGGAAGAACCTGAAG	817		
DB	433	CGSCACAGCATTCCTGAAGAGGATTCCTCTGGCTATGAATCCTGCCACAGAGTATTCGAAG	374		
QY	818	AGACTCTGTTGTCGAGCAGTGCAAATTTGACCAAAGATCTCAACCCCAATCTCTCGGAAAGG	877		
DB	373	ATCAGCAGGATGATATGAAGCCGGAAATGATACCGATGATCTAAGGAAGACGGACAC	314		
QY	878	AAGCTGATCAGGTTGTCGACAGTGCATTTTGACCAAGATTCGAAACCAAGTCTCTGGGC	937		
DB	313	ACCTTCACCACTTCGGCCTGATGGTCTCATTCAGAAAGATCCCGAAGATATAGATGGGC	254		
QY	938	AGAAATGTGACGAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAA	988		

	Query Match	1.2%;	Score 37.4;	DB 4;	Length 1407;
	Best Local Similarity	43.6%;	Pred. No. 0.63;		
	Matches 167;	Conservative	0;	Mismatches 216;	Indels 0;
	Gaps	0;			
Qy	489	ATCTCAAGGGAAGATGCTGATAGATCAACTCTTCACAAAGAGTGTGCAAGAAAGCCAATGA	548		
Db	558	AGCTAAACGTAAAGCTGAGACTGAAGCAAGACTTAAACCGTGAAGCTCAAAAAGCTGAAA	617		
Qy	549	CTCCAATGCAATGCGCCTTTCTGGCAAGAATGAGCTGCTGAGGCCAATACTGATTCACC	608		
Db	618	CGCTAAACTTTCAGCAACAACAAGAGCTAAACGAAAGGCCGACAGATGCTTAAGCCAA	677		
Qy	609	AATGTAAGATTTTCAAGGGCCAGCCCAAAATATGATGTGGCAGCAAAATGCTCTCTGAGGA	668		
Db	678	ACMGAAAAGCTGCCGAGACGCCAAAGCTTAAAGCTGAATCAAGTCTTAAGCTTAAGCA	737		
Qy	669	CAACACTTCCTGTGATGTTGGGGCTTTTACCTGAAGTTCGCCAGATTTACATGGCACATAGA	728		
Db	738	ACAAGCAGCTGATTAACCGCAACCGTAAAGCCGAGGCCGATGCGAAAGCTTAAACAGCAAAA	797		
Qy	729	AGTAAATGCTGCAGATCAACCTCCATCCACTCCAAACTTTTCTGAAGTGGTCTCTCAAAAG	788		
Db	798	AGCAGCTGAAGATGCTTAAACGTTAAAGCTGAGGCTGATGCAAAAGCCCAAGCAGCAAAAGC	857		
Qy	789	AAATGAAGATGAAATGGAAGAAACTGAAGAGACTCTTGTGTGCTGACAGTGCAAATTTGAC	848		
Db	858	AGCTGAAGATGCTAAGCGCAGAGCTGAGGCCGATGCGAAAGCTTAACAGCAAAAAGCGGC	917		
Qy	849	CAAAGATCTTAACCCCAATGCTTG	871		
Db	918	TGAAGATGCTTAAACGCAAAAGCTG	940		

```

RESULT 12
US-09-489-039A-6914
; Sequence 6914, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Braton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6914
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6914

```



```

? LOCATION: (5046)..(6307)
? OTHER INFORMATION: exon 6
? NAME/KEY: intron
? LOCATION: (6308)..(6447)
? NAME/KEY: exon
? LOCATION: (6448)..(8065)
? OTHER INFORMATION: exon 7
? NAME/KEY: intron
? LOCATION: (8066)..(8300)
? NAME/KEY: exon
? LOCATION: (8301)..(8648)
? OTHER INFORMATION: exon 8
? US-09-415-946-2

```

Query Match	1.2%	Score 36.8	DB 4	Length 8648
Best Local Similarity	50.6%	Pred. No. 3.7		
Matches	89	Conservative	0	Mismatches 87
				Indels 0
				Gaps 0
QY	1724	AGGTTTCGACGTGAGCATGATATCCAAATTTATGCTCGACCTTCATGAGCAGAGTCATC	1783	
DBb	6446	AGGATGCTTCCTGGTCCAGATGAAGAGGGGAGACTGTTATGGTCCAGAGACATCATGGAG	6505	
QY	1784	CCAGACAGAAAAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATG	1843	
DBb	6506	CACCAAGAGGCCAAGTCCAGNTAGAGGAGACTACGACTGAAGAGCAAAACAAGATG	6565	
QY	1844	ACATCCCATGGATATTTGTAAGTCTAGCTTAATAAAACACAGCATGAGAGGCGAGCTT	1999	
DBb	6566	ATATTCCAAATGAGAGATGTGGAGCTCATGGCCAAAAAACAGTATCAGAGAGGTGCTTT	6621	

```

RESULT 14
US-09-415-946-1
; Sequence 1, Application US/09415946
; Patent No. 6376751
; GENERAL INFORMATION:
; APPLICANT: Sung, Z. Renee
; APPLICANT: Aubert, Dominique
; APPLICANT: Chen, Lingling
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids that Control Reproductive Development in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 018941-000110US
; CURRENT APPLICATION NUMBER: US/09/415,946
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,696
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 17341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: genomic DNA (Ecotype Columbia) from CD82 clone
; NAME/KEY: promoter
; LOCATION: (1)..(3201)
; NAME/KEY: CDS
; LOCATION: join(4241..4335, 4448..4623, 4704..4823, 4903..4956,
; OTHER INFORMATION: EMBRYONIC FLOWER 1 (EMFL)
; NAME/KEY: exon
; LOCATION: (3202)..(3265)
; OTHER INFORMATION: exon 1
; NAME/KEY: intron
; LOCATION: (3266)..(4159)
; NAME/KEY: exon
; LOCATION: (4160)..(4335)
; OTHER INFORMATION: exon 2
; NAME/KEY: intron
; LOCATION: (4336)..(4447)
; NAME/KEY: exon
; LOCATION: (4448)..(4623)
; OTHER INFORMATION: exon 3

```

NAME/KEY: intron
LOCATION: (4624)..(4703)
NAME/KEY: exon
LOCATION: (4704)..(4823)
OTHER INFORMATION: exon 4
NAME/KEY: intron
LOCATION: (4824)..(4902)
NAME/KEY: exon
LOCATION: (4903)..(4956)
OTHER INFORMATION: exon 5
NAME/KEY: intron
LOCATION: (4957)..(5045)
NAME/KEY: exon
LOCATION: (5046)..(6307)
OTHER INFORMATION: exon 6
NAME/KEY: intron
LOCATION: (6308)..(6447)
NAME/KEY: exon
LOCATION: (6448)..(8065)
OTHER INFORMATION: exon 7
NAME/KEY: intron
LOCATION: (8066)..(8300)
NAME/KEY: exon
LOCATION: (8301)..(8648)
OTHER INFORMATION: exon 8
US-09-415-946-1

Query Match 1.2%; Score 36.8; DB 4; Length 17341;
Best Local Similarity 50.6%; Pred. No. 6.2;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 1724 AGTTTCTCCGCTGAGCATGATATCCAAATTATGTCGACCTTCATGAGCAGTCTAC 1783
DB 6446 AGGATGCTTCTGGTGAGATAGAAAGGGGAGACTGTTATGGTCCAAAGACATCATGGAG 6505
QY 1784 CCAAGAGAAAAGAGACAAAACCTTGAAGTGACTCGTGAAGAAACAGACCATCATAGATG 1843
DB 6506 CACCAAGAGCCAAAGTCACGATAGAGAGAGACTAGCACTGAAGAGCAAAACAACGATG 6565
QY 1844 ACATCCCATGATATGTTGAAGTCTGCTAGCTTAAGAAACAGCATGAGAGCGAGCTT 1899
DB 6566 ATATCCATGGAGATAGTGGAGCTCATGCGCAAAACCAAGTACGAGAGGTGTCTT 6621

RESULT 15
US-08-676-967-2
Sequence 2, Application US/08676967
Patent No. 5747317
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSES: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,967
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCS96-055
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-967-2
Query Match 1.1%; Score 36.4; DB 1; Length 2277;
Best Local Similarity 29.7%; Pred. No. 1.8;
Matches 116; Conservative 66; Mismatches 206; Indels 3; Gaps 2;
QY 668 ACAACACTCTGTTGATGTTGGGGCTTTTACCTGAGTTCCCGAGATTACATGCGACATAG 727
DB 391 AARACNGTNTTYGNCARTTYGCGCGTNYTNGARGTNAAYATHCCNMGNAARCCNGAY 450
QY 728 AAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACCTTTCTGAAGTGGTCCCTCAAAA 787
DB 451 CGNAARATGCGNGGNTTYGGNTTGTNCARTTYAARAAAYTNTYNGARGCGNGGNAARCN 510
QY 788 GAATGAAGATGAAAATGAAAACCTGAAGAGAC-TCTTTGTTGCTGAGCAGTGCATTTTG 846
DB 511 YTNARCGNATGAAYATGAARGATHAARGNMGNACNGTNGCTNGAYTGGCGCTN 570
QY 847 ACCAAGATCTTAACCCCAATGCTCTGGAAGGAACGTGATCAGTTGCTGAGCAGTGCAT 906
DB 571 GCAARGAYAAATYAARGAYACNCARWSNGTNSNGCNATHGGNGARGAARWSNCAY 630
QY 907 TTGACCAAGATCCGAACACAGTGTCTGGGCGAAGATCTGAGCAGATCTGCAATCGACCA 966
DB 631 GARWSNARCAICARGARWSGTNABARARGGNMGNARGARGARGAYATGARGAR 690
QY 967 TGTGAAGAAGTTGTTCTCAAAAAGAGCTTCAAAATCTAAGAGGAAGACCGATAAGAGTTG 1026
DB 691 GARGAARAYGAGV--AYGAYGAYGAYGAYGAYGARGARGAYGCGTNTTYGAYGAYG 748
QY 1027 ATGAAGAGCAGCGCACAGCAAGAACGCA 1057
DB 749 ARGAYGARGARGARGAARAAAYATHGARNNA 779

Search completed: May 17, 2004, 01:46:55
Job time : 224 secs

GenCore version S.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2004, 23:34:20 ; Search time 1288 Seconds
(without alignments)
11172.479 Million cell updates/sec

Title: US-09-721-114-1_COPY_655_3825

Perfect score: 3171
Sequence: 1 atggagatgtgcagtaga.....acaggtatattgattaccgc 3171

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	3171	100.0	3896	9 US-09-828-068-1	Sequence 1, Appli
2	3171	100.0	4310	15 US-10-447-135-1	Sequence 1, Appli
3	2573.2	81.1	9455	15 US-10-447-135-3	Sequence 3, Appli
4	168.8	5.3	871	13 US-10-425-114-4042	Sequence 4042, Ap
5	73.4	2.3	279	9 US-09-294-0938-2759	Sequence 2759, Ap
6	42.2	1.3	319630	16 US-10-398-221-7	Sequence 7, Appli
7	42.2	1.3	3011208	16 US-10-398-221-2058	Sequence 2058, Ap
8	41	1.3	1260	13 US-10-282-122A-34801	Sequence 34801, A
9	40.6	1.3	5451	13 US-10-188-186-59	Sequence 59, Appl
10	40.6	1.3	5761	13 US-10-302-172-23	Sequence 23, Appl
11	40	1.3	412	10 US-09-918-995-22363	Sequence 22363, A
12	39.8	1.3	15122	15 US-10-311-455-748	Sequence 748, Appl
13	39.8	1.3	3673778	15 US-10-312-841-2	Sequence 2, Appli
14	39.2	1.2	1251	13 US-10-425-114-813	Sequence 813, App

Sequence 114, App
Sequence 477, Appl
Sequence 1, Appli
Sequence 2106, Ap
Sequence 147, App
Sequence 451, App
Sequence 19524, A
Sequence 25517, A
Sequence 557, App
Sequence 8851, App
Sequence 324274,
Sequence 324274,
Sequence 146, App
Sequence 237, App
Sequence 621, App
Sequence 621, App
Sequence 215, App
Sequence 99680, A
Sequence 99680, A
Sequence 170, App
Sequence 170, App
Sequence 24, Appl
Sequence 6666, Ap
Sequence 18143, A
Sequence 45, Appl
Sequence 45, Appl
Sequence 27, Appl
Sequence 82, Appl
Sequence 82, Appl
Sequence 82, Appl
Sequence 82, Appl

9 US-09-745-763-114
15 US-10-311-455-477
16 US-10-085-117-1
12 US-10-152-319A-2106
12 US-10-041-018-147
13 US-10-087-192-451
2 US-09-864-761-19524
261
9 US-09-864-761-25517
9 US-09-864-761-557
9 US-09-864-761-8851
13 US-10-027-632-324274
16 US-10-027-632-324274
9 US-09-770-445-146
16 US-10-310-154-237
9 US-09-938-842A-621
11 US-09-938-842A-621
17 US-10-149-310-215
13 US-10-027-632-99680
16 US-10-027-632-99680
9 US-09-925-302-170
13 US-09-925-302-170
16 US-10-439-388-24
10 US-09-764-891-6666
9 US-09-894-844-45
13 US-10-647-089-45
16 US-10-388-902-45
15 US-10-197-666A-27
15 US-10-024-298A-82
15 US-10-042-211A-82
17 US-10-617-217A-82

ALIGNMENTS

RESULT 1

US-09-828-068-1
; Sequence 1, Application US/09828068
; Patent No. US20020157137A1
; GENERAL INFORMATION:
; APPLICANT: Moon, Yong-Hwan
; APPLICANT: Chen, Lingling
; APPLICANT: Sung, Zimay R.
; TITLE OF INVENTION: NUCLEIC ACIDS THAT CONTROL REPRODUCTIVE DEVELOPMENT IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 018941-001400US
; CURRENT APPLICATION NUMBER: US/09/828,068
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 3896
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (241)...(3411)
US-09-828-068-1

Query Match 100.0%; Score 3171; DB 9; Length 3896;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGATTTGTCAGTAGATCAGGAGGAGCTGCTGTTGTTGGGACGAACTGTATGCTT 60
DB 241 ATGGAGATTTGTCAGTAGATCAGGAGGAGCTGCTGTTGTTGGGACGAACTGTATGCTT 300
QY 61 GCTCGTGTGGAACTGTGCTGTAGCCGACGTTGTTGGAGCTGACACGACGCTGCTCAG 120
DB 301 GCTCGTGTGGAACTGTGCTGTAGCCGACGTTGTTGGAGCTGACACGACGCTGCTCAG 360
QY 121 GATGACGCCGCTGAAGCTGGTGTAGAGAACCGGACAAACCAATGCGAGCATTTCTCC 180

Db	361		GATCAGCGCTGAAGCTGGTGTAGACGAAACCGGCACACCAATGCGAGCATTTCTCC	420
Qy	181		ATAAGAGGTATGTTGCTCTCTTTCAGAAAGGATCCAAATTTCTGCTCTCTATCTCGG	240
Db	421		ATAAGAGGTATGTTGCTCTCTTTCAGAAAGGATCCAAATTTCTGCTCTCTATCTCGG	480
Qy	241		ATTTTCAATGACGAGAAATAATGTATGATGAAACAAAGCTAGTTCAAGCCCATTTCTGTA	300
Db	481		ATTTTCAATGACGAGAAATAATGTATGATGAAACAAAGCTAGTTCAAGCCCATTTCTGTA	540
Qy	301		GCAAGTTTCAAGCATGGATGTTGCTTCAAGTGTCTTGGATAGTTGAAACTTTCAGATAAT	360
Db	541		GCAAGTTTCAAGCATGGATGTTGCTTCAAGTGTCTTGGATAGTTGAAACTTTCAGATAAT	600
Qy	361		GGACAGCACCAAGAATCTTTCCGCAAAAGCAGAAATGGCAAGATGATGTTGCTCCATC	420
Db	601		GGACAGCACCAAGAATCTTTCCGCAAAAGCAGAAATGGCAAGATGATGTTGCTCCATC	660
Qy	421		ACATTTGTTGGAGCATCTTTTGTCCTGCTAGTGTGTTGGTTCCTCCAAAGTGTCTCTAGC	480
Db	661		ACATTTGTTGGAGCATCTTTTGTCCTGCTAGTGTGTTGGTTCCTCCAAAGTGTCTCTAGC	720
Qy	481		ACAAATCATCTCAAGGGAAGATGCTGATAGATCAACTCTTCCAAAGAGTGTCCAGAA	540
Db	721		ACAAATCATCTCAAGGGAAGATGCTGATAGATCAACTCTTCCAAAGAGTGTCCAGAA	780
Qy	541		GGCAATGACTCCAAATGCAATGCGCTTCTGGCAAGAAATGGAGCTGCTGAGGCCAATACT	600
Db	781		GGCAATGACTCCAAATGCAATGCGCTTCTGGCAAGAAATGGAGCTGCTGAGGCCAATACT	840
Qy	601		GATTTCACCAATGAAAGATTTGCAAGGCGCCAGCCCAAATATGATGTGGCGACAAATGTC	660
Db	841		GATTTCACCAATGAAAGATTTGCAAGGCGCCAGCCCAAATATGATGTGGCGACAAATGTC	900
Qy	661		TCTGAGGACAACTCTCTGTTGATGTTGGGCTTTTACCTGAAGTTTCCCAAGATTAATG	720
Db	901		TCTGAGGACAACTCTCTGTTGATGTTGGGCTTTTACCTGAAGTTTCCCAAGATTAATG	960
Qy	721		CACATAGAGTAAATGFTGCAGATCAACCTCCATCCACTCCTCAAACTTTCTGAAGTGGTC	780
Db	961		CACATAGAGTAAATGFTGCAGATCAACCTCCATCCACTCCTCAAACTTTCTGAAGTGGTC	1020
Qy	781		CTCAAAAGAAATGAAGATGAAATGGAAAACTGAAGAGACTCTTGTGCTGAGCAGTGC	840
Db	1021		CTCAAAAGAAATGAAGATGAAATGGAAAACTGAAGAGACTCTTGTGCTGAGCAGTGC	1080
Qy	841		AAHTTGACCAAGATCTCTAACCAATGCTCGGAAGGAACTGATCAGGTTGCTGAGCAG	900
Db	1081		AAHTTGACCAAGATCTCTAACCAATGCTCGGAAGGAACTGATCAGGTTGCTGAGCAG	1140
Qy	901		TGCAATTTGACAAAGATCCGAACCAAGTGTCTGGCAGAAATCTGAGCAGATCTGCAAT	960
Db	1141		TGCAATTTGACAAAGATCCGAACCAAGTGTCTGGCAGAAATCTGAGCAGATCTGCAAT	1200
Qy	961		GAGCCATGTGAAGATTTGTTCTCAAAAGAGCTTCCAAATCTAAGAGAAACGGATTAAG	1020
Db	1201		GAGCCATGTGAAGATTTGTTCTCAAAAGAGCTTCCAAATCTAAGAGAAACGGATTAAG	1260
Qy	1021		AAGTTTGATGAAGAGCAGCAGCAGCAAGAAACGCACTGCCAGGCTGATGTTTCAGAT	1080
Db	1261		AAGTTTGATGAAGAGCAGCAGCAGCAAGAAACGCACTGCCAGGCTGATGTTTCAGAT	1320
Qy	1081		GCAAGCTTTGTCGGAGAAAGCCAAAAAAGTGGCGTTCTATCAGAAATTAATAATGCT	1140
Db	1321		GCAAGCTTTGTCGGAGAAAGCCAAAAAAGTGGCGTTCTATCAGAAATTAATAATGCT	1380
Qy	1141		AAACAGTTGAGGATTTCTAGAGTGAAGAGTTTCACTGTTGAAATGCGCTGATCCCTGT	1200
Db	1381		AAACAGTTGAGGATTTCTAGAGTGAAGAGTTTCACTGTTGAAATGCGCTGATCCCTGT	1440
Qy	1201		GAGGATGATAGAAGTACATCCCGGCTCCGATGGAAGTGAAGCATGAGATATTCCTGTTAGC	1260

Db	1441	GAGGATGATAGAAAGTACCATCCCGGTCCCGATGGGAAGTAGCATGGATATTCCTGTTAGC	1500
Qy	1261	AAACATACAGTGGGAGAAGATGGGTTAAAAATCAAGTAAGAAACAAAGCAAAACGCCAAATAC	1320
Db	1501	AAACATACAGTGGGAGAAGATGGGTTAAAAATCAAGTAAGAAACAAAGCAAAACGCCAAATAC	1560
Qy	1321	TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAAATGGAAAAAAGAAAGA	1380
Db	1561	TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAAATGGAAAAAAGAAAGA	1620
Qy	1381	ACTGGAAGTGTGCATCACACAGTTGCTCAFCACAGCTGGGAAATTTGAGCAACAAAAAAGTG	1440
Db	1621	ACTGGAAGTGTGCATCACACAGTTGCTCATCAGCTGGGAAATTTGAGCAACAAAAAAGTG	1680
Qy	1441	ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATCTGAAAAATGGTCTTGACACA	1500
Db	1681	ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATCTGAAAAATGGTCTTGACACA	1740
Qy	1501	AATATGATGAAGACAGATGTCGTGACAGCATGATCAGAAATCTCCACACAGAGGTGCTCA	1560
Db	1741	AATATGATGAAGACAGATGTCGTGACAGCATGATCAGAAATCTCCACACAGAGGTGCTCA	1800
Qy	1561	TCAAAGGGGAAAAACAGCGGGTTTGAGTAAGGGGAAAAACATTCAGCTGCTAGTACCAAA	1620
Db	1801	TCAAAGGGGAAAAACAGCGGGTTTGAGTAAGGGGAAAAACATTCAGCTGCTAGTACCAAA	1860
Qy	1621	TATGGTGGTGAAGACACACAGGAATGGTCAGACATACATGTACTCAGCGCAGAGATCAA	1680
Db	1861	TATGGTGGTGAAGACACACAGGAATGGTCAGACATACATGTACTCAGCGCAGAGATCAA	1920
Qy	1681	TGCAGATGAAACCGAAAACTCTGTTCTTGAGTCACTCGCGAAAAGTTTTCTCCAGCTGAG	1740
Db	1921	TGCAGATGAAACCGAAAACTCTGTTCTTGAGTCACTCGCGAAAAGTTTTCTCCAGCTGAG	1980
Qy	1741	CATGATATCCAAATTTATGTTCTGACCTTTATGAGCAGAGTCTACCCAGAGAAAGAG	1800
Db	1981	CATGATATCCAAATTTATGTTCTGACCTTTATGAGCAGAGTCTACCCAGAGAAAGAGAG	2040
Qy	1801	CAAAAACTTGAAGTGACTCGTGAANAACAGACCATGATAGATGACATCCCCATGGATATT	1860
Db	2041	CAAAAACTTGAAGTGACTCGTGAANAACAGACCATGATAGATGACATCCCCATGGATATT	2100
Qy	1861	GTTGAACCTGCTAGCTAAAAACAGCATGAGAGCGAGCTTATGACTGAGACTGATTTGTTCT	1920
Db	2101	GTTGAACCTGCTAGCTAAAAACAGCATGAGAGCGAGCTTATGACTGAGACTGATTTGTTCT	2160
Qy	1921	GACATCAACGGTATTCAAATCCCAAGACAACTGCTGATGATGATTTGTTGTAATAGTAGCTGCC	1980
Db	2161	GACATCAACGGTATTCAAATCCCAAGACAACTGCTGATGATGATTTGTTGTAATAGTAGCTGCC	2220
Qy	1981	AAGGATGGTTCAGATTTATGCATCAAGTGTGTTTGACACTAATTTCCCAACAGAAAGTCTTG	2040
Db	2221	AAGGATGGTTCAGATTTATGCATCAAGTGTGTTTGACACTAATTTCCCAACAGAAAGTCTTG	2280
Qy	2041	GCATCCCAAGTACACAGAGAGGTTTACAGGTCATTTGGCAATTCACACACAGAGTCT	2100
Db	2281	GCATCCCAAGTACACAGAGAGGTTTACAGGTCATTTGGCAATTCACACACAGAGTCT	2340
Qy	2101	CCACATCCTCAGAACTTTTCCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA	2160
Db	2341	CCACATCCTCAGAACTTTTCCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA	2400
Qy	2161	ATGGTCACTATTGCTGCAAGCTCACCACTATTTTACATCATCATGATCAGTATATTGCT	2220
Db	2401	ATGGTCACTATTGCTGCAAGCTCACCACTATTTTACATCATCATGATCAGTATATTGCT	2460
Qy	2221	GAAGCAACCAACTGMACTTTGGGCGCGTAAAGACGCAAGAAAGCTAACTGGGAGCAATTTT	2280
Db	2461	GAAGCAACCAACTGMACTTTGGGCGCGTAAAGACGCAAGAAAGCTAACTGGGAGCAATTTT	2520
Qy	2281	AAGGCCACTACAAGAAATTTCTCCAGACGAAACATGTGGTGTCTCAATTTAGACCTGGTATC	2340
Db	2521	AAGGCCACTACAAGAAATTTCTCCAGACGAAACATGTGGTGTCTCAATTTAGACCTGGTATC	2580

Qy	2341	CAAGCAGTTGCATTTGACTTCTTACTCATGTGTCATGGGATCTTTCCAGCAATTAATGACATCTCGC	2400
Db	2581	CAAGCAGTTGACTTTGACTTCTACTCATGTGTCATGGGATCTTTCCAGCAATTAATGACATCTCGC	2540
Qy	2401	CAACCAAGTAATTCGGCCACCTGGACCGCTATGCTGAAAGAGCGGTAAACAGSTCCATGCA	2460
Db	2641	CAACCAAGTAATTCGGCCACCTGGACCGCTATGCTGAAAGAGCGGTAAACAGSTCCATGCA	2700
Qy	2461	AGAAAATTTTCCAAAGCACAAATAGCAACCAATGSAAGCGAGTAAGTTATGTGATCGGAGAAAT	2520
Db	2701	AGAAAATTTTCCAAAGCACAAATAGCAACCAATGSAAGCGAGTAAGTTATGTGATCGGAGAAAT	2760
Qy	2521	GCTGGCAAGTAGTCTTTGTATCTTAAAGATTCATGCTGCGAGCGATCTTCTTGAGAAATG	2580
Db	2761	GCTGGCAAGTAGTCTTTGTATCTTAAAGATTCATGCTGCGAGCGATCTTCTTGAGAAATG	2820
Qy	2581	ATGGATCCATCAACATTAGCAAGCTTCCCAAATATGGAATCTTCTAGCAGGAACCCAGATG	2640
Db	2821	ATGGATCCATCAACATTAGCAAGCTTCCCAAATATGGAATCTTCTAGCAGGAACCCAGATG	2880
Qy	2641	GAGTCTCAACTTCATTAATTTCTAGTAGTGACATAATCAGTACAAAGGATCAACACGACCA	2700
Db	2881	GAGTCTCAACTTCATTAATTTCTAGTAGTGACATAATCAGTACAAAGGATCAACACGACCA	2940
Qy	2701	TCATATGGCAGTAGTAACCTGGAATGGAAGATTCCTATGCAATTTGGAAGACTTTATCACGGCAT	2760
Db	2941	TCATATGGCAGTAGTAACCTGGAATGGAAGATTCCTATGCAATTTGGAAGACTTTATCACGGCAT	3000
Qy	2761	CAGCTGCAATGATCTGCAACAGACCTTTTAGGCCACATCCTAGAGTTGGTGTGCTTGCTGCTC	2820
Db	3001	CAGCTGCAATGATCTGCAACAGACCTTTTAGGCCACATCCTAGAGTTGGTGTGCTTGCTGCTC	3060
Qy	2821	TTGCTGCAGAAAGGAATTTGCAAACTGGTTCGGAGAACTGTGGCCACACAATCTGGTTATTAAG	2880
Db	3061	TTGCTGCAGAAAGGAATTTGCAAACTGGTTCGGAGAACTGTGGCCACACAATCTGGTTATTAAG	3120
Qy	2881	TTAGGAGTGTCAACAGAAATPAAATCCATCCATCAGATGAACAGAAAGGAACATTTTGAAGCC	2940
Db	3121	TTAGGAGTGTCAACAGAAATPAAATCCATCCATCAGATGAACAGAAAGGAACATTTTGAAGCC	3180
Qy	2941	CTGAATTCCTGGAATGTTTTCCAGCAAAAATGGAATGCAATTCAGATTTGGGTTCTGTTAGCTCC	3000
Db	3181	CTGAATTCCTGGAATGTTTTCCAGCAAAAATGGAATGCAATTCAGATTTGGGTTCTGTTAGCTCC	3240
Qy	3001	AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTTGGACCGAGGCAAGGGT	3060
Db	3241	AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTTGGACCGAGGCAAGGGT	3300
Qy	3061	AAAATGGTTTATCCCTTGGATCGGTTTGTGACAGCAGGATATCTGTATAACTAACAAGAAC	3120
Db	3301	AAAATGGTTTATCCCTTGGATCGGTTTGTGACAGCAGGATATCTGTATAACTAACAAGAAC	3360
Qy	3121	CCAGCTGATTTTACTACAATCAGTAAACGATAACGAGTATATGGATTACCGC	3171
Db	3361	CCAGCTGATTTTACTACAATCAGTAAACGATAACGAGTATATGGATTACCGC	3411

RESULT 2
US-10-447-135-1
; Sequence 1, Application US/10447135
; Publication No. US20030199684A1
; GENERAL INFORMATION:
; APPLICANT: Hirochika, Hirohiko
; APPLICANT: Yamazaki, Munee
; APPLICANT: Miyao, Akio
; TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
; FILE REFERENCE: MARF-1 DIV
; CURRENT APPLICATION NUMBER: US/10/447,135
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US 09/721,114
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JAPAN 2000-1149106

/	PRIOR FILING DATE:	2000-05-19					
/	NUMBER OF SEQ ID NOS:	3					
/	SOFTWARE:	Patent In Ver. 2.1					
/	SEQ ID NO 1						
/	LENGTH:	4310					
/	TYPE:	DNA					
/	ORGANISM:	Oryza sativa					
/	FEATURE:						
/	NAME/KEY:	CDS					
/	LOCATION:	(655)..(3928)					
/	US-10-447-135--1						
			Query Match 100.0%; Score 3171; DB 15; Length 4310;				
			Best Local Similarity 100.0%; Pred No. 0;				
			Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATCGAGATTGTTCCAGTAGACATCAGGAGGGAGCTCGTGTGTTTGCGAGAGCAACTGTATGCTT	60				
Db	655	ATCGAGATTGTTCCAGTAGACATCAGGAGGGAGCTCGTGTGTTTGCGAGAGCAACTGTATGCTT	714				
Qy	61	GCTCGTGTGGAACTGGTGCTGTAGCGCCAGTCGTCGAGCTGCACAGCGCCTCGTCAG	120				
Db	715	GCTCGTGTGGAACTGGTGCTGTAGCGCCAGTCGTCGAGCTGCACAGCGCCTCGTCAG	774				
Qy	121	GATCGACGCCCTGAAGCTGGTGTAGACGAAACCGGCAACAACCAATGCGAGCATTTCTTC	180				
Db	775	GATCGACGCCCTGAAGCTGGTGTAGACGAAACCGGCAACAACCAATGCGAGCATTTCTTC	834				
Qy	181	ATAAGGGGTATCTTCTCTCTTTCAGAGAAGATCCAAAATTCTGCTCTCTATCTCGG	240				
Db	835	ATAAGGGGTATCTTCTCTCTTTCAGAGAAGATCCAAAATTCTGCTCTCTATCTCGG	894				
Qy	241	ATTTTCCATGACCACAGAAAAAATGTGTAGAACCAAAGCTAGTTCAAGCCCATTTCGTGA	300				
Db	895	ATTTTCCATGACCACAGAAAAAATGTGTAGAACCAAAGCTAGTTCAAGCCCATTTCGTGA	954				
Qy	301	GCAAGTTTCGACGTGGATTCCTCGAAGTGTCTTGATAGTTGAAACTTCAGATAT	360				
Db	955	GCAAGTTTCGACGTGGATTCCTCGAAGTGTCTTGATAGTTGAAACTTCAGATAT	1014				
Qy	361	GGAAACAGCAACAAGAACTCTTCCCGCAAGACGAATGGCACAGTGAATGTTCTCCATC	420				
Db	1015	GGAAACAGCAACAAGAACTCTTCCCGCAAGACGAATGGCACAGTGAATGTTCTCCATC	1074				
Qy	421	ACATTTGTTCCGAGCACTTTTGTGCTGCTGATGPTGGTTCCAAAAGTGCTCTCTAGC	480				
Db	1075	ACATTTGTTCCGAGCACTTTTGTGCTGCTGATGPTGGTTCCAAAAGTGCTCTCTAGC	1134				
Qy	481	ACCAATCATCTCAAGGAGAGATGCTGTAGATCAACTCTTCCAAGAGTGTGCAAGAA	540				
Db	1135	ACCAATCATCTCAAGGAGAGATGCTGTAGATCAACTCTTCCAAGAGTGTGCAAGAA	1194				
Qy	541	GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGATGGAGCTGCTGAGGCCAATACT	600				
Db	1195	GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGATGGAGCTGCTGAGGCCAATACT	1254				
Qy	601	GATTCCCAATGAAAGATTGCAAGGCCAGCCCAAAATTATGATCTGCGACCAATGTC	660				
Db	1255	GATTCCCAATGAAAGATTGCAAGGCCAGCCCAAAATTATGATCTGCGACCAATGTC	1314				
Qy	661	TCTGAGCAACAACCTTCTGTGTATGTTGGGGCTTTTACCTGAACTTCCCAGATTACATGG	720				
Db	1315	TCTGAGCAACAACCTTCTGTGTATGTTGGGGCTTTTACCTGAACTTCCCAGATTACATGG	1374				
Qy	721	CACATGAGTAATAGTGTGAGATCAACCTCGATCCACTCCAAAATCTTCTGAAGTGGTC	780				
Db	1375	CACATGAGTAATAGTGTGAGATCAACCTCGATCCACTCCAAAATCTTCTGAAGTGGTC	1434				
Qy	781	CTCAAAAGAAATCAAGATGAAAAATGAAAAATCTGAAGAGACTCTTGTGCTGAGCAGTGC	840				
Db	1435	CTCAAAAGAAATCAAGATGAAAAATGAAAAATCTGAAGAGACTCTTGTGCTGAGCAGTGC	1494				
Qy	841	AATTTGCCAAAGATVCTTAACCCAAATGTCTGGAAGAGAACCTGATCAGTGTGCTGAGCAG	900				

1495 Db |||||GACCAAGATCTTAACCAATGCTGGAAGGACGATGATCAGGTTCTGAGCAG 1554
901 QY TCGAATTTGACCAAGATCGGAACCAAGTGTCTGGGACAGAAATGTGAGCAGATCTGCAAT 960
1555 Db TCGAATTTGACCAAGATCGGAACCAAGTGTCTGGGACAGAAATGTGAGCAGATCTGCAAT 1614
961 QY GAGCCATGTGAAGAGTGTCTCTCAAAAGAAAGTCTCAAAATCTTAAGAGGAAGACGATTAAG 1020
1615 Db GAGCCATGTGAAGAGTGTCTCTCAAAAGAAAGTCTCAAAATCTTAAGAGGAAGACGATTAAG 1674
1021 QY AAGTTGATGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1080
1675 Db AAGTTGATGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1734
1081 QY GCAAGCTTTGTCGGAGGAAGCAAAAGAGTGGGCTCTCTATCAGAAATTAATAATGCT 1140
1735 Db GCAAGCTTTGTCGGAGGAAGCAAAAGAGTGGGCTCTCTATCAGAAATTAATAATGCT 1794
1141 QY AACCAAGTTGAGGATTTCTAGAAGTGAAGAGTTCATGCTGAAATTCGCGCTGATCCCTGT 1200
1795 Db AACCAAGTTGAGGATTTCTAGAAGTGAAGAGTTCATGCTGAAATTCGCGCTGATCCCTGT 1854
1201 QY GAGGATGATGAAGTACCAATCCCGTCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260
1855 Db GAGGATGATGAAGTACCAATCCCGTCCGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1914
1261 QY AACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGCAAGCAAAACGCAAAATAC 1320
1915 Db AACCATACAGTGGGAGAGATGGGTTAAATCAAGTGAAGCAAGCAAAACGCAAAATAC 1974
1321 QY TCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1975 Db TCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2034
1381 QY ACTGGAAGTGTGCATCACACAGTGTGCTCATCAGCTGGGAAATTCGAGCAACAAAAGAGTG 1440
2035 Db ACTGGAAGTGTGCATCACACAGTGTGCTCATCAGCTGGGAAATTCGAGCAACAAAAGAGTG 2094
1441 QY ACACCCACTGCGAGTACTCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
2095 Db ACACCCACTGCGAGTACTCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2154
1501 QY AATATGATTAAGACAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
2155 Db AATATGATTAAGACAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2214
1561 QY TCAAGGGGAAACAGCGGTTTGATTAAGGGGAAACACATTCAGCTGCTAGTACCAAA 1620
2215 Db TCAAGGGGAAACAGCGGTTTGATTAAGGGGAAACACATTCAGCTGCTAGTACCAAA 2274
1621 QY TATGTTGTTGAAGACACAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
2275 Db TATGTTGTTGAAGACACAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2334
1681 QY TGCAGATGGAACCGAAACCTGTGTTCTGAGTCACTCGGCAAGGTTTCTCCAGCTGAG 1740
2335 Db TGCAGATGGAACCGAAACCTGTGTTCTGAGTCACTCGGCAAGGTTTCTCCAGCTGAG 2394
1741 QY CATGATATCCAAATATGTTCTGACCTTCATGAGCAGGTTACCCAGAGGAAAGAG 1800
2395 Db CATGATATCCAAATATGTTCTGACCTTCATGAGCAGGTTACCCAGAGGAAAGAG 2454
1801 QY CAAAACTTGAAGTGAATCTCGTCAAAACAGACCATGATGATGATGATGATGATGATGATGATGATGAT 1860
2455 Db CAAAACTTGAAGTGAATCTCGTCAAAACAGACCATGATGATGATGATGATGATGATGATGATGATGAT 2514
1861 QY GTTGAATGCTAGCTAAAAACAGGATGAGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1920
2515 Db GTTGAATGCTAGCTAAAAACAGGATGAGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGAT 2574
1921 QY GACATCAACCGTATTCAATCCAGACAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980

2575 Db GACATCAACCGTATTCAATCCAGACAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634
1981 QY AAGGATCGTTGAT 2040
2635 Db AAGGATCGTTGAT 2694
2041 QY GCATCCCAAGATACACAGAGGAGTTACAGGCTCATTTGGCATTTGACACACACAGAGTCT 2100
2695 Db GCATCCCAAGATACACAGAGGAGTTACAGGCTCATTTGGCATTTGACACACACAGAGTCT 2754
2101 QY CCACATCTCAGAACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2755 Db CCACATCTCAGAACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2814
2161 QY ATGTCCTACTATTTGTCGAAGCTCACCACTATTTTTCATCATCATGATGATGATGATGATGATGATGATGAT 2220
2815 Db ATGTCCTACTATTTGTCGAAGCTCACCACTATTTTTCATCATCATGATGATGATGATGATGATGATGATGAT 2874
2221 QY GAAGCAACCACTGAACTTTGGGCGGTAAAGACCAAGAGCTAAAGCTGGGACCAATTT 2280
2875 Db GAAGCAACCACTGAACTTTGGGCGGTAAAGACCAAGAGCTAAAGCTGGGACCAATTT 2934
2281 QY AAGGCCACTACAGAAATTTCTCAGCAGCAACATGCTGCTCAATTTAGACCTGGTATC 2340
2935 Db AAGGCCACTACAGAAATTTCTCAGCAGCAACATGCTGCTCAATTTAGACCTGGTATC 2994
2341 QY CAGCAGTGTGAT 2400
2995 Db CAGCAGTGTGAT 3054
2401 QY CAACCAAGTAAATTTGGGCACTGGACCGCTATGCTGAAAGAGCGGTAAACCCAGGTCCTATGCA 2460
3055 Db CAACCAAGTAAATTTGGGCACTGGACCGCTATGCTGAAAGAGCGGTAAACCCAGGTCCTATGCA 3114
2461 QY AGAAATTTTCCAGCAACATAGCAACCAATGGAAGCGAGTAAATTTGATGATGATGATGATGATGATGATGAT 2520
3115 Db AGAAATTTTCCAGCAACATAGCAACCAATGGAAGCGAGTAAATTTGATGATGATGATGATGATGATGATGAT 3174
2521 QY GCTGGAAGAGTGTGAT 2580
3175 Db GCTGGAAGAGTGTGAT 3234
2581 QY ATGATCCATCAACATTAGCAAGCTTCCCAACTATGGAATCTTACGAGGAAACAGATG 2640
3235 Db ATGATCCATCAACATTAGCAAGCTTCCCAACTATGGAATCTTACGAGGAAACAGATG 3294
2641 QY GAGTCTCAACTTCAATTTCTCAGTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
3295 Db GAGTCTCAACTTCAATTTCTCAGTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3354
2701 QY TCATATGCGAGTAACTGATGGAAGATTTCCATTGACATTTGGAAGATTTTACGAGGATTTATCAACGAGAT 2760
3355 Db TCATATGCGAGTAACTGATGGAAGATTTCCATTGACATTTGGAAGATTTTACGAGGATTTATCAACGAGAT 3414
2761 QY CAGCTGAT 2820
3415 Db CAGCTGAT 3474
2821 QY TTGCTGCAAGAGGAAATTTGCAAACTGGTGGAGAACTGTGGCACAACATTTCTGTTTATAAG 2880
3475 Db TTGCTGCAAGAGGAAATTTGCAAACTGGTGGAGAACTGTGGCACAACATTTCTGTTTATAAG 3534
2881 QY TTAGAGTGTCAA CAGGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
3535 Db TTAGAGTGTCAA CAGGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3594
2941 QY CTGAATTTCTGAAATTTTTCAGCAAAATGGAATGCAATTTGAGTTGGGTTCTGTTAGCTTCC 3000
3595 Db CTGAATTTCTGAAATTTTTCAGCAAAATGGAATGCAATTTGAGTTGGGTTCTGTTAGCTTCC 3654
3001 QY AGTGCAGATTTTATCAGCAGGAAACAGCATAGCTCAATCTTGGACAGAGGAGGAT 3060
3655 Db AGTGCAGATTTTATCAGCAGGAAACAGCATAGCTCAATCTTGGACAGAGGAGGAT 3714

QY 3061 AAAATGTTTCATCCCTGGATCGTTTGTGAGCAGGATATCTGTATAACTAACAGAAC 3120
Db |||||
QY 3715 AAAATGTTTCATCCCTGGATCGTTTGTGAGCAGGATATCTGTATAACTAACAGAAC 3774
Db |||||
QY 3121 CCAGCTGATTTTACTACAATCAGTAACGATAACAGATATGATTAACCGC 3171
Db |||||
QY 3775 CCAGCTGATTTTACTACAATCAGTAACGATAACAGATATGATTAACCGC 3825
Db |||||
RESULT 3
US-10-447-135-3
; Publication 3, Application US/10447135
; Publication No. US20030199684A1
; GENERAL INFORMATION:
; APPLICANT: Hirochika, Hirohiko
; APPLICANT: Yamazaki, Muneo
; APPLICANT: Miyao, Akio
; TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
; FILE REFERENCE: NAF-1 DIV
; CURRENT APPLICATION NUMBER: US/10/447,135
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US 09/721,114
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JAPAN 2000-149106
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 9455
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-10-447-135-3
Query Match 81.1%; Score 2573.2; DB 15; Length 9455;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 3003; Conservative 0; Mismatches 23; Indels 396; Gaps 2;
QY 146 ACGAACCGGCACACACCAATGGAGCAATTCCTCCATAGAGGGTATGTCCTCTTC 205
Db |||||
QY 5447 ATGCATCTTTATATGATTAATCTTATTTCTTGTACAGAGGGTATGTCCTCTTC 5506
Db |||||
QY 206 AGAAGAGGATCCAAATCTCTCTCTCTCTCTCTCTCGATTTCCATGACCAAGAAATGTC 265
Db |||||
QY 5507 AGAAGAGGATCCAAATCTCTCTCTCTCTCTCTCGATTTCCATGACCAAGAAATGTC 5566
Db |||||
QY 266 ATGAACAAAGCTAGTTCGAAGCCCAATTTCTGTAGCAAAAGTTTCGAAGATGGATGCT 325
Db |||||
QY 5567 ATGAACAAAGCTAGTTCGAAGCCCAATTTCTGTAGCAAAAGTTTCGAAGATGGATGCT 5626
Db |||||
QY 326 CGAAGTCTTGGATAGTTCGAAGCTTCAGATATGGAACAGCAGCAGCAAGTCTTCOCG 385
Db |||||
QY 5627 CGAAGTCTTGGATAGTTCGAAGCTTCAGATATGGAACAGCAGCAGCAAGTCTTCOCG 5686
Db |||||
QY 386 CAAAGCAGAAATGGCACAAGTGTGTTGCTCCATCACAATTTGTTCCGAGCACTTTTGTGC 445
Db |||||
QY 5687 CAAAGCAGAAATGGCACAAGTGTGTTGCTCCATCACAATTTGTTCCGAGCACTTTTGTGC 5746
Db |||||
QY 446 CTGCTAGTGTGTTTCCCAAAAGTGTCTCTAGCAGCAGCAATCATCTCAAGGGAAGATG 505
Db |||||
QY 5747 CTGCTAGTGTGTTTCCCAAAAGTGTCTCTAGCAGCAGCAATCATCTCAAGGGAAGATG 5806
Db |||||
QY 506 CTGATAGTCAACTCTTCCAAAGAGTGTGCAAGAGGCAATGCACTCCAAATGCAATGCGC 565
Db |||||
QY 5807 CTGATAGTCAACTCTTCCAAAGAGTGTGCAAGAGGCAATGCACTCCAAATGCAATGCGC 5866
Db |||||
QY 566 CTTCGCGAAGATGGAGCTGTGAGGCCAATCTGATTCACCAATGAA----- 614
Db |||||
QY 5867 CTTCGCGAAGATGGAGCTGTGAGGCCAATCTGATTCACCAATGAAAGTATGGTAG 5926
Db |||||
QY 615 ----- 614
Db |||||
QY 5927 ATGTAGAGCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGATAACTAATGT 5986
Db |||||

QY 615 -----AGATTTCCAGGGCCAGCCCAAAATTTATGATGTCGCGCAAAATGT 659
Db |||||
QY 5987 TTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTTATGATGTCGCGCAAAATGT 6046
Db |||||
QY 660 CTCTGAGGCAACACACTTCTGTTTGATGTTGGGGCTTTTACCTGAAGTTCCTCCAGATTAATG 719
Db |||||
QY 6047 CTCTGAGGCAACACACTTCTGTTTGATGTTGGGGCTTTTACCTGAAGTTCCTCCAGATTAATG 6106
Db |||||
QY 720 GCACATAGAGTAATGTTGTCAGATCAACCTCCATCCACTCCAAACTTTCTGAACTGCT 779
Db |||||
QY 6107 GCACATAGAGTAATGTTGTCAGATCAACCTCCATCCACTCCAAACTTTCTGAACTGCT 6166
Db |||||
QY 780 CCTCAAAAGAAATGAAGATGAAATGGAATACTGAAGAGACTCTTGTGTCGAGCAGTG 839
Db |||||
QY 6167 CCTCAAAAGAAATGAAGATGAAATGGAATACTGAAGAGACTCTTGTGTCGAGCAGTG 6226
Db |||||
QY 840 CAAATTTGACCAAGATCTTAAACCCCAATGTCCTGGAAGGAACGTGATCAGTTCCTGAGCA 899
Db |||||
QY 6227 CAAATTTGACCAAGATCTTAAACCCCAATGTCCTGGAAGGAACGTGATCAGTTCCTGAGCA 6286
Db |||||
QY 900 GTGCAATTTGACCAAGATCTTAAACCCCAATGTCCTGGAAGGAACGTGATCAGTTCCTGAGCA 959
Db |||||
QY 6287 GTGCAATTTGACCAAGATCTTAAACCCCAATGTCCTGGAAGGAACGTGATCAGTTCCTGAGCA 6346
Db |||||
QY 960 TGAGCCATGTCGAAGATGTTTCTCAAAAGAAAGTCTCCAAATCTTAAGAGGAACGATTA 1019
Db |||||
QY 6347 TGAGCCATGTCGAAGATGTTTCTCAAAAGAAAGTCTCCAAATCTTAAGAGGAACGATTA 6406
Db |||||
QY 1020 GAAGTTGATGAAGAGCAGCAGCAGCAAGAAACGACTGCTGCCAGGCTGATGTTTTCAGA 1079
Db |||||
QY 6407 GAAGTTGATGAAGAGCAGCAGCAGCAAGAAACGACTGCTGCCAGGCTGATGTTTTCAGA 6466
Db |||||
QY 1080 TGCAAAAGCTTTCTCGAGAAAGCCCAAAAGTGGCGGCTTCTATCAGAAATTTATAATGC 1139
Db |||||
QY 6467 TGCAAAAGCTTTCTCGAGAAAGCCCAAAAGTGGCGGCTTCTATCAGAAATTTATAATGC 6526
Db |||||
QY 1140 TAAACAGGTTGAGGATTTCTAGAGTGCAGAGTTCATGCTGGAATTCGCGTGCATCCCTG 1199
Db |||||
QY 6527 TAAACAGGTTGAGGATTTCTAGAGTGCAGAGTTCATGCTGGAATTCGCGTGCATCCCTG 6586
Db |||||
QY 1200 TCAGATGATAGAGTACCATCCCGTCCCGATGGAAGTAAAGTATGATATTCCTGTTAG 1259
Db |||||
QY 6587 TGAGGATGATAGAGTACCATCCCGTCCCGATGGAAGTAAAGTATGATATTCCTGTTAG 6646
Db |||||
QY 1260 CAAACATACAGTGGGGAAGATGGGTTAAATCAAGTAAAGAAAGCAAGCAAAACGCAATA 1319
Db |||||
QY 6647 CAAACATACAGTGGGGAAGATGGGTTAAATCAAGTAAAGAAAGCAAGCAAAACGCAATA 6706
Db |||||
QY 1320 CTCTGATGTTGTAGATGATGATCATCTTATGAACTGGCTGAATGGAATAAGAAAAAG 1379
Db |||||
QY 6707 CTCTGATGTTGTAGATGATGATCATCTTATGAACTGGCTGAATGGAATAAGAAAAAG 6766
Db |||||
QY 1380 AACTGGAAGTGTGCATCAACAGTTCCTCATCCAGTGGGAATTTGAGCAACAAAAAGT 1439
Db |||||
QY 6767 AACTGGAAGTGTGCATCAACAGTTCCTCATCCAGTGGGAATTTGAGCAACAAAAAGT 6826
Db |||||
QY 1440 GACACCCCTGCGAGTACTCAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Db |||||
QY 6827 GACACCCCTGCGAGTACTCAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 6886
Db |||||
QY 1500 AATATGCAATAGACAGATGCTGTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 1559
Db |||||
QY 6887 AATATGCAATAGACAGATGCTGTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 6946
Db |||||
QY 1560 ATCAAGGGGAAAAACAGCGGTTTGTAGTAAAGGGAACACATTCGCTGCTAGTACCAA 1619
Db |||||
QY 6947 ATCAAGGGGAAAAACAGCGGTTTGTAGTAAAGGGAACACATTCGCTGCTAGTACCAA 7006
Db |||||
QY 1620 ATATGTTGTGAAAGCAGCAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 1679
Db |||||
QY 7007 ATATGTTGTGAAAGCAGCAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 7066
Db |||||

FEATURE:
US-10-425-114-4042

Query Match
Best Local Similarity 5.3%; Score 168.8; DB 13; Length 871;
Matches 299; Conservative 0; Mismatches 132; Indels 21; Gaps 3;

QY 2720 ATGGAAGATTCCATTGCAATTCGAAGACTTATCAGCGCATCA---GCTGCATGATCTGC 2776
DB 105 ATGGAAGCCCAACCGCTGAGACTTGTCTCGCGCTCATTCNCCAGCAAGACTTGC 164

QY 2777 ACAGACCTTTAGCCGCCACATCTAGAGTTGGTGTGCTTGGCTCCTTCTGCAGAGGAAA 2836
DB 165 GCAGGCTTTAGGCTTACGCTGACCTGCGGCTGCTCGGTTCTGCTGAGCGAGGA 224

QY 2837 TTGCAAACTGGTTCGAGAACTGTGGCACAACTCTGGTTATAGTTAGGAG 2887
DB 225 TCGCAAACTGGTCTCGGAA-TGCGGGCCGCGCTGTGGGTANAGACTAGGTG 274

RESULT 6
US-10-398-221-7/c
; Sequence 7, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 319630
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-7

Query Match 1.3%; Score 42.2; DB 16; Length 319630;
Best Local Similarity 47.0%; Pred. No. 32;
Matches 131; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 574 AAGAATGGAGCTGCTGAGGCAATATCTGATTCACCAATGAAGATTGCAAGGCCAGCC 633
DB 289425 AACAAAAAGACATGNAATTTATGGCCAATATGCCACAGAGAGTTTCTGTGTGCG 289366

QY 634 CAAATTTATGATGGGAGCAAAATGCTCTGAGGACAACTTCTGTGTGTTGGGCT 693
DB 289365 GATAAACTAAAGTGAAGGTTATGTTTCAAGCAGCAAAACCACTTAGTTATGAGGAAAT 289306

QY 694 TTACCTGAAGTCCCGAGATTACATGCGCATAGAGTAATGTCGAGATCAACCTCCA 753
DB 289305 ATTATGATGACTTCAAAATCACTTTGCAAAAAGTCTGTTTGTGGCTGAGCTGCA 289246

QY 754 TCCACTCCAAAACCTTTCTGAAGTGGTCTCTCAAAAGAAATGAAGATGAAAATGAAAAC 813
DB 289245 TCTGGCGAGAAATTTTAAAGATTTAATCGCAACGATGAAGGTTCTCATTTATTAGGC 289186

QY 814 GAAGAGACTCTTCTGCTGAGCAGTGCNATTTGACCAA 852
DB 289185 GAAGTGGCTCTAGTTCAGATCCATCCCTATTTCCTCA 289147

RESULT 7
US-10-398-221-2058
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications

FEATURE:
US-10-425-114-4042

Query Match
Best Local Similarity 5.3%; Score 168.8; DB 13; Length 871;
Matches 299; Conservative 0; Mismatches 132; Indels 21; Gaps 3;

QY 2720 ATGGAAGATTCCATTGCAATTCGAAGACTTATCAGCGCATCA---GCTGCATGATCTGC 2776
DB 105 ATGGAAGCCCAACCGCTGAGACTTGTCTCGCGCTCATTCNCCAGCAAGACTTGC 164

QY 2777 ACAGACCTTTAGCCGCCACATCTAGAGTTGGTGTGCTTGGCTCCTTCTGCAGAGGAAA 2836
DB 165 GCAGGCTTTAGGCTTACGCTGACCTGCGGCTGCTCGGTTCTGCTGAGCGAGGA 224

QY 2837 TTGCAAACTGGTTCGAGAACTGTGGCACAACTCTGGTTATAGTTAGGAGTGTCAACAG 2896
DB 225 TCGCAAACTGGTCTCGGAACTGCGGCGCGAGTCTGGGTACAGACTAGTGTGTAAG 284

QY 2897 GAATTAACATGATCAGATGAACAGAAAGAACATTTTGAAGCCTGAAATTTGGAATGT 2956
DB 285 GGAAGGCGCGCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGA----- 332

QY 2957 TTTGAGCAAAATGGAATGATTCAGTTGGTTCCTGTAGCT-----CCAGTGCAGATT 3010
DB 333 GCTCGGAGGATGGAGACCTCGAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGT 392

QY 3011 TTTTATCAGGAGGAGCAAGCATAGCTCAATCTTGACCCAGAGGCAAGGTTAAATGGTTC 3070
DB 393 ACCGGTTTCCATGTTATGTTAGTACGGCTTCAGCCCTTCAGCAGTGGCAATGGGAGGACCTTC 452

QY 3071 ATCCCTTCGATCGTTTCTGAGACAGGATATCTGTATTAACAAAGAACCCAGCTGATT 3130
DB 453 ATCCGTTTGGTAAGCTCTGAGGAGGATATCTGTGTGACTTAAACAGAAACCCAGCCGATT 512

QY 3131 TTACTACATCAGTAACGATAACAGTATATG 3162
DB 513 TCACGTGTAATAGTACAGAACAGGATACATG 544

RESULT 5
US-09-294-093B-2759
; Sequence 2759, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 2759
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700346612H1
; NAME/KEY: unsure
; LOCATION: 28, 123, 150, 179, 191, 264
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-2759

Query Match 2.3%; Score 73.4; DB 9; Length 279;
Best Local Similarity 71.3%; Pred. No. 8.2e-11;
Matches 122; Conservative 0; Mismatches 45; Indels 4; Gaps 2;

FILE REFERENCE: 344 702 - US
CURRENT APPLICATION NUMBER: US/10/398,221
PRIOR FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: FR 00/12 697
PRIOR FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 4025
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2058
LENGTH: 3011208
TYPE: DNA
ORGANISM: Listeria innocua
US-10-398-221-2058

Query Match 1.3%; Score 42.2; DB 16; Length 3011208;
Best Local Similarity 47.0%; Pred. No. 1.6e+02;
Matches 131; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 574 AGAATGGAGCTCTGAGGCAATCTGATTCACCAATGAAGATTCGAAGGGCCAGCC 633
DB 1827998 AACAAAAAGGACATGAATTTATGGCCAAATATGCCAACAGAGAGCTTTCTGTGTGGG 1828057
QY 634 CAAATATATGCTGCGCAGCAAAATGCTCTGAGGACACACTTCTGTGATGTGGGGCT 693
DB 1828058 GATRAACTAAAGTGAAGGTATGTTTCAAGCAGAACCACTTAGTTATGCAGGAAT 1828117
QY 694 TTACCTGAAGTCCCGAGATTACATGGCACAATAGAAATGATGGTCAGATCAACTCCA 753
DB 1828118 ATTATTGATGACTTCAAAATCACTTTCCGAAAAGGTCGTATTGTTGGCGTTGAAAGTCCA 1828177
QY 754 TCCACTCCAAAAGCTTCTGAGGAGGTCCTCAAAAGAAATGAAGATCAAAATGAAAAACT 813
DB 1828178 TCTGGCGAGAAATTTTAAAGATTTTAAATCGCAAGATGAAGTTCTCATTTATTAGGC 1828237
QY 814 GAAGACACTTTGTTGCTGAGCAGTCCAAATTTGACCAA 852
DB 1828238 GAAGTGGCTTAGTTCAGATCCATCCCTATTTCCAA 1828276

RESULT 8
US-10-282-122A-34801
Sequence 34801, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34801
LENGTH: 1260
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-10-282-122A-34801

Query Match 1.3%; Score 41; DB 13; Length 1260;
Best Local Similarity 49.8%; Pred. No. 1.2;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 473 CTCCTAGCACCAATCATCTCAAGGGAAGATGCTGATGATCAACTCTTCCAAAGAGTG 532
DB 1052 CTATAAGTGAAGAAAGCAATCGAACGTAAACAGGAGCGCGTGATTACGTTCAATTATAG 1111
QY 533 TGCAGAGGCAATCACTCCAAATGCAATGCGCTTCTGGCAAGATCGAGCTGCTGAGG 592
DB 1112 AAGAGCAATTAATGCAATCATGATGATGATGATGATGATGATGATGATGATGATG 1171
QY 593 CCAATACTGATTCACCAATGAAGATTCGAAGGCGCAGCCCAAAATATGATGATGCGCAG 652
DB 1172 TCATTACTGAACAAACATTAATGAAGAAATGAGCCTGAATTATGATGATGATGATG 1231
QY 653 CAAATGCTCTGAGGACCAACTCTGTT 681
DB 1232 ATCTAATTAATGAAGCAAGACATCTGCT 1260

RESULT 9
US-10-188-186-59
Sequence 59, Application US/10183186
Publication No. US20040029789A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-397C
CURRENT APPLICATION NUMBER: US/10/188,186
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/303046
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/360814
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/303828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/323380
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/361133
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/304016
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/304502
PRIOR FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom
SEQ ID NO 59
LENGTH: 5451

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(5146)
US-10-188-186-59

Query Match      1.3%; Score 40.6; DB 13; Length 5451;
Best Local Similarity 48.9%; Pred. No. 4.7;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 1716 CTCGGCAAGGTTTCTCCAGCTGAGCATGATATCCAAATTTATGTTGACCTTCATGAGCA 1775
Db 2235 CTTGGTAGATCTTACTCAGAGAAAGCAAAAAGTTAGTCTGTAATATGAAAGAA 2294

Qy 1776 GAGTCTACCAAGAAAGAGCAAACTTGAAGTCACTCGTGAAAGACAGACCAT 1835
Db 2295 ACTCCAAGCCATGAACAAGAACTGCGAGACITTCAGCGAGCTCAAAAAGACATGCAAG 2354

Qy 1836 GATAGATGATATCCCATGATATTTGTAAGTCTAGCTTAAAGACAGCATGAGAGCA 1895
Db 2355 GTTGCTTAAATAATCAGTCTCAGTATGAAAGCAATTTGAAGAAATTCGACGAGGATGTGAT 2414

Qy 1896 GCTTATGATGAGACTGATTTGTTCTGACATCAACCGTATTCAA 1938
Db 2415 GGAATGAAAAAACAAGGTTCCGCTTAATGAAACAAATGAAA 2457
```

```

RESULT 10
US-10-302-172-23
; Sequence 23, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
; FILE REFERENCE: 803-1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 23
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(5148)
US-10-302-172-23
```

```

Query Match      1.3%; Score 40.6; DB 13; Length 5761;
Best Local Similarity 48.9%; Pred. No. 4.9;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 1716 CTCGGCAAGGTTTCTCCAGCTGAGCATGATATCCAAATTTATGTTGACCTTCATGAGCA 1775
Db 2235 CTTGGTAGATCTTACTCAGAGAAAGCAAAAAGTTAGTCTGTAATATGAAAGAA 2294

Qy 1776 GAGTCTACCAAGAAAGAGCAAACTTGAAGTCACTCGTGAAAGACAGACCAT 1835
Db 2295 ACTCCAAGCCATGAACAAGAACTGCGAGACITTCAGCGAGCTCAAAAAGACATGCAAG 2354

Qy 1836 GATAGATGATATCCCATGATATTTGTAAGTCTAGCTTAAAGACAGCATGAGAGCA 1895
Db 2355 GTTGCTTAAATAATCAGTCTCAGTATGAAAGCAATTTGAAGAAATTCGACGAGGATGTGAT 2414
```

```

Qy 1896 GCTTATGATGAGACTGATTTGTTCTGACATCAACCGTATTCAA 1938
Db 2415 GGAATGAAAAAACAAGGTTCCGCTTAATGAAACAAATGAAA 2457

RESULT 11
US-09-918-995-22363
; Sequence 22363, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22363
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-22363
```

```

Query Match      1.3%; Score 40; DB 10; Length 412;
Best Local Similarity 48.3%; Pred. No. 1.1;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 1247 ATATTCTCTTTAGCAACCATACAGTGGGAGAGATGGGTTAAATCAAGTAAAGACAAGA 1306
Db 168 ACATCCAGGAACCAAGGCTTAGAGAGGGAACATGGACCTTTGGAGGAGCTACCTAC 227

Qy 1307 CAAAAGCGAAATACTCTGATGTTGTAGATGATGATGATCATCTATGAACTGGCTGAATG 1366
Db 228 GAGGTGCCAGATGCCCTCGAGTGGGACAGAGATCAAGGCTTGTGAGAGCTGGATG 287

Qy 1367 GAAAAAGAAAAAGAACTGGAAGTGTGATCAACAGCTGCTCATCCAGCTGGGAATTCGA 1426
Db 288 GAAAAATTTAAAGACAAGGCTGTTACATGTTGGGATGCTGCTGCTGCTGCTGGGTTGGG 347

Qy 1427 GCACACAAAAAGTGAACCCACTGCGAGTACTCAGCATGATGAGATGA 1478
Db 348 ATAGCTGGAGCCTTGTTCATCTTGGAGCCCTCTACAGCATTAAGGTTATGA 399
```

```

RESULT 12
US-10-311-455-748/c
; Sequence 748, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 748
; LENGTH: 15122
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-748

Query Match
Best Local Similarity 1.3%; Score 39.8; DB 15; Length 15122;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1763 ACCTTCATGAGCAGCTCTACCAAGAGAAAGCAAAACTTGAAGTGAATCGTG 1822
Db AACTAACGACCAAACTTAACCAAAATAATAAAAAAATAAATAAATAAATA 6636

QY 1823 AAAACAGACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1882
Db AATAAAAAACCAAAATATATTAATCACTACCAAAATAAATAAATAAATA 6576

QY 1883 AGCATGAGAGGAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1942
Db ACTAAACTATAAAAAATCACTCTAATACCGTATTAATCACTCAACATAA 6516

QY 1943 AGCAAC 1949
Db AATAAAC 6509

RESULT 13
US-10-312-841-2
Sequence 2, Application US/10312841
Publication No. US20030186277A1
GENERAL INFORMATION:
APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
FILE REFERENCE: E01/1208/WO
CURRENT APPLICATION NUMBER: US/10/312,841
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
LENGTH: 3673778
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: (379615)
US-10-312-841-2

Query Match
Best Local Similarity 1.3%; Score 39.8; DB 15; Length 3673778;
Matches 80; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1278 AGATGGTTAAATCAAGTAAGAACAAAGCAAAACGCAAAATCTCTGATGTTGTAGATGA 1337
Db AGATTGTTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAATTTAGATAGTAGTTTA 2884356

QY 1338 TGGATCATCACTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1397
Db AGGGTTAAAAAATTTTATATATAGTTGTTAAAGTGAGAGAGGATTTGGTATTATGAATAA 2884416

QY 1398 CACAGTTGCTCATCCAGCTGGGAATTT 1424
Db TATAGTTGATATATATAGTTATATAT 2884443

RESULT 14
US-10-425-114-813
Sequence 813, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 813
LENGTH: 1251
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700106469_FLI
US-10-425-114-813

Query Match
Best Local Similarity 1.2%; Score 39.2; DB 13; Length 1251;
Matches 107; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 578 ATGAGCTGCTGAGGCCAATCTCTCTGAGGACAACTTCTGTTGATTTGGGCTTTAC 637
Db ATGTTGAAGTTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 784

QY 638 ATTATGATGTGGCAGCAAAATGCTCTCTGAGGACAACTTCTGTTGATTTGGGCTTTAC 697
Db TTCTGCTGTGGACTCAAGCGGAGATATAATCAGAGATGCGCATGATGAGGCTATAT 844

QY 698 CTGAAGTTCCCGCAGATTACATGCGCACATAGAGTAATGTTGTCAGATCAACCTCCATCCA 757
Db TTGGTGTACCTACAACTTCATTTGTTGATGTTGAATTTGGTGTGACGAGCTTAAACCA 904

QY 758 CTCCAAACTTTCTGAAAGTGTCTCTCAAAAGAAATGAAGA 797
Db ATCATGAGACGGGAAAGTTGAGACACACAGATCTGATGA 944

RESULT 15
US-09-745-763-114
Sequence 114, Application US/09745763
Patent No. US20020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Marberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridge Park Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284


```
TELEPAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 114:
US-09-745-763-114

Query Match      1.2%; Score 39.2; DB 9; Length 2522;
Best Local Similarity 47.4%; Pred. No. 7;
Matches 110; Conservative 2; Mismatches 120; Indels 0; Gaps 0;

Qy 1247 ATATTCTGTAGCAACCATACAGTGGGAGAGATGGGTTAAATCAAGTAAGACAAGA 1306
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 2102 ACATCCCGAAGACCGCTTAGAGGAGGGAAACATGGACCTGTGGAGGGAGCTACCTAC 2161
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 1307 CAAAGCGCAANTACTCTGATGTTGTAGATGATGATCATCACCTTATGAAGTGGCTGAATG 1366
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 2162 CAGGTCCAGATGCCCTCGAGTGGGACAGAGATCAAGGCTGGTGAGAGCTGGATG 2221
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 1367 GAAAGAGAAAGAAATGGAAGTGTGCATCACACAGTGTCTCATCCAGCTGGGAATTTGA 1426
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 2222 GAAAGATTAAAGACAAAGGCTGTTACATCTCTGGGATGCTGTGCTGTAGGGGTTGGG 2281
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 1427 GCAACAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAGAAATGA 1478
Db      |||      |||      |||      |||      |||      |||      |||      |||
Qy 2282 ATAGCTGGAGCCTTGTTCATCTTTGGAGGCCCTACAGCATTAAGGTTAIGA 2333
Db      |||      |||      |||      |||      |||      |||      |||      |||
```

Search completed: May 17, 2004, 05:12:31
Job time : 1332 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2004, 19:11:26 ; Search time 7740 seconds
(without alignments)
12234.234 Million cell updates/sec

Title: US-09-721-114-1_COPY_655_3825

Perfect score: 3171
Sequence: 1 atggagattgttcagtaga.....acgagtatatggattaccgc 3171

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estm.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hc.*
- 9: gb_esti.*
- 10: gb_est2.*
- 11: gb_hc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pin.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gsa1.*
- 29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	355.6	11.2	1025	28	BZ795917 PUFHM45TD
2	351.2	11.1	888	28	CC355521 PUFHPL93TB
3	268.2	8.5	523	12	BJ246359 BJ246359
4	237.4	7.5	885	29	CG091207 PUFJK24TD

5	236	7.4	469	29	AG213697	Oryza sat
6	229.6	7.2	593	13	BQ244810	Tab15035B
7	205.8	6.5	999	28	BZ795914	PUFHM45TB
8	205.2	6.5	556	13	BU974305	HB27110r
9	203	6.5	787	29	CG345663	CGHLM77TH
10	178.6	5.6	469	9	AL818058	AL818058
11	171.2	5.4	491	29	AG209720	Oryza sat
12	168.8	5.3	641	28	BZ411511	BZ411511
13	167.2	5.3	849	28	BZ411504	CGACC43TC
14	152	4.8	766	29	CC603916	CGWHM55TV
15	139.4	4.4	635	28	BH880296	ht52d11.b
16	129	4.1	822	29	CG365016	OGHAD70TV
17	124.4	3.9	753	28	CC160109	CC160109
18	124.4	3.9	995	29	CG451574	OGAD34TV
19	122.8	3.9	842	28	BZ705032	PUCSC90TD
20	120.8	3.8	703	28	BZ776688	1104h10.b
21	119.6	3.8	688	28	BZ776689	1104h10.g
22	118.8	3.7	970	29	CG032005	PUFTA25TD
23	114.2	3.6	653	28	BZ311518	ic58h08.b
24	114.2	3.6	870	29	CG091452	PUFHF81TB
25	111.2	3.5	903	28	BZ704364	BZ704364
26	103.6	3.3	793	28	BZ423675	1d52a11.g
27	103.2	3.3	782	29	CC824592	ZMMBB016
28	98.6	3.1	701	28	BZ374744	ie31d12.g
29	91.2	2.9	590	28	CC160110	CC160110
30	91	2.9	562	28	CC04201	PUDDK15TD
31	91	2.9	792	28	BZ422300	BZ422300
32	91	2.9	809	28	BZ652757	OGAM082TM
33	83.8	2.6	905	29	CG365008	OGHAD70TH
34	82.4	2.6	664	13	BQ242025	Tab15035B
35	80.2	2.5	600	29	CC603904	CGWHM55TH
36	78.2	2.5	183	12	BZ252288	BJ252288
37	78.2	2.5	389	29	CC991021	ZUABP49TH
38	77.4	2.4	628	29	CC744988	ZMMBB012
39	77.4	2.4	734	28	BZ326387	BZ326387
40	76.6	2.4	632	13	BQ161601	BQ161601
41	76	2.4	587	28	BZ359680	BZ359680
42	75	2.4	321	9	AI691520	606020C11
43	74	2.3	707	29	CG837058	ZMMBB021
44	72.6	2.3	562	28	BZ652753	OGAM082TC
45	67.2	2.1	373	29	CG831985	ZMMBB010

ALIGNMENTS

RESULT 1
BZ795917/c
LOCUS
DEFINITION
BZ795917
BZ795917
BZ795917.1
GI:28996773
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and Bennett, J.
TITLE
JOURNAL
COMMENT
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFHM45TB
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF

BZ795917 1025 bp DNA linear GSS 17-MAR-2003
PUFHM45TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBT320H17,
genomic survey sequence.

Accession BZ795917
Version BZ795917.1
Keywords GSS.
Source Zea mays
Organism Zea mays
Reference
Authors
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and Bennett, J.
Title
Journal
Comment
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFHM45TB
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF

Class: sheared ends.

FEATURES
source 1..1025
/location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 11.2%; Score 355.6; DB 28; Length 1025;
Best Local Similarity 67.0%; Pred. NO. 5.5e-76;
Matches 619; Conservative 0; Mismatches 254; Indels 51; Gaps 6;
2099 CTCCACATCTCTGAGACTTTTCACTCTCAGTACAGGACGACGACACATTTCCGGATGAGG 2158
888 CTCTCTCTGAGACTTTTGGAGTGATCCCAAGACACGACGACACATTTCCAGGGGATG 829
2159 AAATGGTCACTATGCTGCAAGCTCACCACATTTTTCACATCATGATGATGATGATG 2218
828 GAGGGGTCACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 775
2219 CTGAAGCACCACCTGACATTTGGGCGGTGAGGACGACGACGACGACGACGACGACG 2278
774 CTGAAGTGGCTGCTGAGGCTGAGGAAATACCGGGGAAAGAGTAAATGTTGGATCTT 715
2279 TTAAAGCCACTACAAGAAATTTCCAGCAGCAACATGCTGCTCAATTTAGACCTGGTA 2338
714 TCAGACAGCTTCAAGAAATTTCCAAACATCGTCATATGTTTCAATTCGGAACAGGC 655
2339 TCAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2398
654 TTCAAGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 595
2399 GCCAACCAAGTAAATTTGGCCACTGACCGCTATGCTGAAAGAGGGTTAACAGGTCATG 2458
594 ACCAGCAGTAAATTTGGCTGTAGA-----TCAGC 565
2459 CAAGAAATTTTCAAGCACAATGAGCAACCAATGGAAGCGAGTAAATGATGAGGAA 2518
564 CGAAGAGTGTTCAGCAGCAACATGCTGATGAGGCTGATGATGATGATGATGATG 505
2519 ATGCTGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578
504 TTCTTGACAGTCAAGCAAACTCTGGATTCACATATACAGCAGTATGATGATGATG 445
2579 TGATGATCCATCAACATGAGCAAGCTTCCCAACTATGGAATTTCTAGCAGGAAACAGA 2638
444 TGATGACTCATCAACAGCTCCAGACTTTCACAACTATCAAGAGCTAACAGGCGCAGA 385
2639 TGGAGTCTCAACTTCA-----ATTCAGTATGAC-----ATATCAGTACAGGAT 2689
384 TGAACATCAACCAAACTCTGGATTCACATATACAGCAGTATGATGATGATGATG 325
2690 CAACCAAGCAGTATGAGGAGTAACTGATGGAAGAGTTCATTTGACATTCGAGAGCT 2749
324 CACTGAGCAGTATGAGGAGCAGATTAATGGAAGGTTCCACCGACATGCAAGACT 265
2750 TATCAGGAGTATG-----CTGATGATGAGCAGAGCTTTTACGCCAGATCTTAGAGT 2806
264 TATCTGGGCTGAGTCCAGCAAACTGACAGAGGCTTTTACGCCCTTCTCTCGGGTGG 205
2807 GTGCTGTGGCTCTGCTGACAGGAAATTCGAACTGCTGAGAGAGTGTGG---CA 2863
204 GCGTGTGTTGTTGCTGACAGGAGTATCGCAACTGGTATGAACTTGGGGCCGC 145
2864 CACAACTGTGTTAAGTATGAGGAGTGTCAACAGGAATAACATGCAATCAGATGAGAA 2923
144 CACAGTCTGAGTACAGGCTGGGGCTTTCTAAAGGGAGCGGATCGCTTGTATGAAACAG 85

2924 AGGAACATTTTGAAGCCCTGAATTTGGAATTTTTCAGCAAAATGGAATGCAATTCAGT 2983
84 ATGAAACTACGAGACCTTGAACCTCAGGAGTGTTCACAGCAGGATGGAATGCCCTTCAAT 25
2984 TGGGTTCTGTTAGCTCCAGTGCAG 3007
24 TGGGTTCTGTTAGCTCTCTGCG 1

RESULT 2
CC355521 888 bp DNA linear GSS 16-MAY-2003
LOCUS PUHPL93TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa504018,
DEFINITION genomic survey sequence.
ACCESSION CC355521
VERSION CC355521.1 GI:30824921
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 888)
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Benetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHPL93TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source 1..888
/location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTa504018"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 11.1%; Score 351.2; DB 28; Length 888;
Best Local Similarity 67.9%; Pred. NO. 6.3e-75;
Matches 602; Conservative 0; Mismatches 233; Indels 51; Gaps 6;
2298 TTCTCCAGCAGCAGCATGCTGCTCAATTTAGACCTGATCCAGCAGTGTGATTCGAC 2357
1 TTCACCAACATCGTCATATGTTTTCATTCGGAACAGGCTTCAGAGTTTCAATCATC 60
2358 TTCTACTCATGTCATGGGATCTCCAGCAATTTATGCAATTCGCAACCAAGTAATTCGCC 2417
61 TCCAGTCACTGTTATGAGCTTCTAGTAATATGATGATGATGATGATGATGATGATG 120
2418 ACTGACCGCTATGCTGGAAGCGGTTAACAGGTCATGCAAGAAATTTTCCAGCAG 2477
121 TATAGT-----CAGCCGAGAGTGTTCAGGAC 150
2478 AATAGCAACCTGGAAGCGAGTAAATTTATGATGCGGAGAAATGCTGGACAGTATGCTT 2537
151 AACATCGACTATGAGGCTGTTAGTTTATGATGATGATGATGATGATGATGATGATG 210
2538 GTATCTTAAGATTCATGCTCGACGCTCTTCTGAGATGATGATGATGATGATGATGAT 2597
211 CTACCAAAAGAGGCTATGCTGCAACATCTTCTGAGTTGATGAGTCTATCAACAGC 270

Class: sheared ends.

FEATURES
source 1..1025
/location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 11.2%; Score 355.6; DB 28; Length 1025;
Best Local Similarity 67.0%; Pred. NO. 5.5e-76;
Matches 619; Conservative 0; Mismatches 254; Indels 51; Gaps 6;
2099 CTCCACATCTCTGAGACTTTTCACTCTCAGTACAGGACGACGACACATTTCCGGATGAGG 2158
888 CTCTCTCTGAGACTTTTGGAGTGATCCCAAGACACGACGACACATTTCCAGGGGATG 829
2159 AAATGGTCACTATGCTGCAAGCTCACCACATTTTTCACATCATGATGATGATGATG 2218
828 GAGGGGTCACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 775
2219 CTGAAGCACCACCTGACATTTGGGCGGTGAGGACGACGACGACGACGACGACGACG 2278
774 CTGAAGTGGCTGCTGAGGCTGAGGAAATACCGGGGAAAGAGTAAATGTTGGATCTT 715
2279 TTAAAGCCACTACAAGAAATTTCCAGCAGCAACATGCTGCTCAATTTAGACCTGGTA 2338
714 TCAGACAGCTTCAAGAAATTTCCAAACATCGTCATATGTTTCAATTCGGAACAGGC 655
2339 TCAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2398
654 TTCAAGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 595
2399 GCCAACCAAGTAAATTTGGCCACTGACCGCTATGCTGAAAGAGGGTTAACAGGTCATG 2458
594 ACCAGCAGTAAATTTGGCTGTAGA-----TCAGC 565
2459 CAAGAAATTTTCAAGCACAATGAGCAACCAATGGAAGCGAGTAAATGATGAGGAA 2518
564 CGAAGAGTGTTCAGCAGCAACATGCTGATGAGGCTGATGATGATGATGATGATG 505
2519 ATGCTGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578
504 TTCTTGACAGTCAAGCAAACTCTGGATTCACATATACAGCAGTATGATGATGATG 445
2579 TGATGATCCATCAACATGAGCAAGCTTCCCAACTATGGAATTTCTAGCAGGAAACAGA 2638
444 TGATGACTCATCAACAGCTCCAGACTTTCACAACTATCAAGAGCTAACAGGCGCAGA 385
2639 TGGAGTCTCAACTTCA-----ATTCAGTATGAC-----ATATCAGTACAGGAT 2689
384 TGAACATCAACCAAACTCTGGATTCACATATACAGCAGTATGATGATGATGATG 325
2690 CAACCAAGCAGTATGAGGAGTAACTGATGGAAGAGTTCATTTGACATTCGAGAGCT 2749
324 CACTGAGCAGTATGAGGAGCAGATTAATGGAAGGTTCCACCGACATGCAAGACT 265
2750 TATCAGGAGTATG-----CTGATGATGAGCAGAGCTTTTACGCCAGATCTTAGAGT 2806
264 TATCTGGGCTGAGTCCAGCAAACTGACAGAGGCTTTTACGCCCTTCTCTCGGGTGG 205
2807 GTGCTGTGGCTCTGCTGACAGGAAATTCGAACTGCTGAGAGAGTGTGG---CA 2863
204 GCGTGTGTTGTTGCTGACAGGAGTATCGCAACTGGTATGAACTTGGGGCCGC 145
2864 CACAACTGTGTTAAGTATGAGGAGTGTCAACAGGAATAACATGCAATCAGATGAGAA 2923
144 CACAGTCTGAGTACAGGCTGGGGCTTTCTAAAGGGAGCGGATCGCTTGTATGAAACAG 85

ORIGIN		/dev_stage="Peekes' scale 10.5.1.1" /clone_lib="y. Ogihara unpublished cDNA library, Wh_f"			
		Query Match	8.5%; Score 268.2; DB 12; Length 523;		
		Best Local Similarity	72.9%; Pred No. 1.3e-54;		
		Matches 345; Conservative	0; Mismatches 128; Indels	0; Gaps	0;
Qy	2121	GTCTACTCAGGAACGACGACACATTTTCGGATGGAAGAAATGGTCATATTGCTGCAAG	2180		
Db	50	GTCTACCCAGGAACACACAGGCACATTCGGGCATCGGAGAAATAGTCATATTGCTGCAAC	109		
Qy	2181	CTCACCACTATTTTCACATCATGATGATCAGTATATTGCTGAGCACCACTGMAACATTTG	2240		
Db	110	CTCACCACTATTTATCACCAGAACGAGCCAGCCTCTTGCTGAGCAACACCTTGAGCGCTG	169		
Qy	2241	GGGCGTGAAGGACGCAAGAACTAACTGGGGAGCAATTTAAGGCCACTACAAAGAAATTC	2300		
Db	170	GAGCCATATGGAGCAAGAGTCGATGTGGAACTTTTCAGGCACCTCCAGGGGATTT	229		
Qy	2301	TCCAGCAGCAACATGGTGTGCTCAATTTAGACCTGGTATCCAAAGCAGTTGACTTGCATTC	2360		
Db	230	ATCAACCAACAACAGGTGTGCTCAGTTCAGACCTAGCATTGACACGGTTGATTTAACTTA	289		
Qy	2361	TACTCATGTCTATGGGATCTTCAGCAATATGTCATCTGCGCAACCGATTAATTCGGCCACT	2420		
Db	290	TACTGATGTGGCGGAGCTAATCGTTATTATCCCACTGCGCAGGCAGTAATTTGCGACCT	349		
Qy	2421	GGACCGCTATGCTGAAGAGAGCGGTTAAACAGGTCCATGCAAGAAATTTTCCAAAGCAAT	2480		
Db	350	TGACCACTATACAAATAGACGATTTAATCTAGTCAGGCAAGAGATTTTCCAAAGTTCACT	409		
Qy	2481	AGCAACCATGGAAGCGAGTAAGTTATGTTGATCGGAGAAATGCTGGCAAGATGATCTTGTA	2540		
Db	410	GTCACCAATGGAAGCTGGTGAATCTGTGTGATCGAAGAAATGTTGACATTCAGGTTTTTA	469		
Qy	2541	TCCTNAGATTCATGCTCGGACGCATCTTCTGAGATGATGGATCCATCAA	2593		
Db	470	TCCAAAGAAACCATGCTGTTACTACAGCCCTGAGATTTGACCCGAGTCAACCA	522		

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 469
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="NDe047"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="T10466T"
/clone_lib="PCR product directly amplified from rice
genomic DNA"
note="The 3' end of retrotransposon Tos17 was found
immediately upstream of this sequence."
1
misc_feature
/note="This insertion point is base 1732 in AF326768."
ORIGIN
Query Match 7.4%; Score 236; DB 29; Length 469;
Best Local Similarity 100.0%; Pred. No. 9.7e-47;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1492 CTTGACACAAATATCATGAAGACAGATGCTGTCTCAGCATGTATCAGAAATCTCCACACAG 1551
Db 1 CTTGACACAAATATCATGAAGACAGATGCTGTCTCAGCATGTATCAGAAATCTCCACACAG 60
QY 1552 AGGTCTCATCAAGGGGAAACACCGGTTTGAGTAAAGGGGAAACACATTCAGCTGCT 1611
Db 61 AGGTCTCATCAAGGGGAAACACCGGTTTGAGTAAAGGGGAAACACATTCAGCTGCT 120
QY 1612 AGTACCAATATGGTGGTGAAGACACCAAGAAATGTCAGAAATCATACATGTACTCAGCGCA 1671
Db 121 AGTACCAATATGGTGGTGAAGACACCAAGAAATGTCAGAAATCATACATGTACTCAGCGCA 180
QY 1672 GAAGATCAATGTCAGATGAACACCGAAACTCTGTTCTGAGTCATCGCAAGGT 1727
Db 181 GAAGATCAATGTCAGATGAACACCGAAACTCTGTTCTGAGTCATCGCAAGGT 236
RESULT 6
BQ244810
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
BQ244810
TaEI5 Triticum aestivum cDNA clone TaEI5035B12R, mRNA
sequence.
BQ244810.1 GI:20440686
EST
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 593)
Cloutier, S.
Wheat functional genomics - Glenlea developing seeds cDNA libraries
Unpublished (2002)

COM:VNT

Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Datoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.gc.ca

was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 035 row: B column: 12
Seq primer: M13 Reverse.
location/Qualifiers
1..593
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TaB15035B12P"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/clone_host="E. coli DH10B"
/clone_lib="TaB15"
/note="Vector: pGWC-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"

FEATURES
source

ORIGIN	Query Match	7.2%	Score 229.6;	DB 13;	Length 593;
	Best Local Similarity	74.0%;	Pred. No. 3.9e-45;		
	Matches 305;	Conservative 0;	Mismatches 104;	Indels 3;	Gaps 1;
QY	2297	ATTCTCCAGCAGCAACATGTGGTCTCAATTTAGACCTGTGTATCCAGCAGCAGTTGACTTGA	2356		
DB					
DB	2	ATTTATCAACCAACACAGGTGTGCTCAGTTCAGACCTAGCATTTGACACGGTTGATTAA	61		
QY	2357	CTTCTACTCATGTCAATGGGATCTTCCAGCAATTTATGTCATCTCGCCACACAGTAATTCGCG	2416		
DB					
DB	62	CTTATACTGATGGCGGGAGCTAATCGTTATTATCCACTCGCCACGACGTAATTTCGA	121		
QY	2417	CAC TGGACCGGTATCTGTAAGAGAGCGGTTAAACAGGTCCTATGCMAGAGAAATTTTCCAGCA	2476		
DB					
DB	122	CAC TTGACCACTATACAANTAGAGCAGCTTAAC T CAGTCCAGGCCAAGAAGTTTTCACAGTT	181		
QY	2477	CAATAGCAACCATGTGAAGCCGAGTAAGTTATGTGTATCGAGAGAAATGCTGGACCAAGTAGTCT	2536		
DB					
DB	182	CAGTGTCAACCATGTGAGCTGTTAATCTGTGTGATGTGAAGAAATGTTGGACATTCAGTT	241		
QY	2537	TGTATCCTTAAGAATTCATCGCTGGGAGCGAATCTTCTGTGAGATGATGGATCCCATCAACAT	2596		
DB					
DB	242	TTTATCCAAGAGAAACCATGCGCTGCTACTCAGCCCGCTTGAGATTGACCGAGTCACCAATGT	301		
QY	2597	TAGCAAGCTTCCCAAGCTATGTGGAACTTCTTAGCAGGAACAGATGGAGTCTCAACTTCATA	2656		
DB					
DB	302	TAGCCAGTTT---CAACTATGAAGGTCTTAGCAGGAACAGATGGAAATTTCAACTTCGGA	358		
QY	2657	ATTCTCAGTATGCACATATCAATCAGTACAAAGATCAACACAGCAATCATATGG	2708		
DB					
DB	359	ATTCACTATGCACAGGATCAGTCCATCGGATCAGCCAGCAGCACACCGTATGG	410		

RESULT 7	BZ795914	999 bp	DNA	linear	GSS 17-MAR-2003
LOCUS	BZ795914				
DEFINITION	PUFM45TB_ZM_0_6_1_0_KB_Zea mays genomic clone ZMWBta320H17, genomic survey sequence.				
ACCESSION	BZ795914				
VERSION	BZ795914.1	GI:28936767			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.				

```

REFERENCE
AUTHORS      1 (bases 1 to 999)
              Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Other GSs: PUPFM45TD
              Contact: Cathy Whitelaw

TIGR
7912 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source       1..999
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="873"
              /db_xref="taxon:4577"
              /clone="ZMMB7A320H17"
              /clone_lib="ZM_0.6_1.0 kb"
              /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"

```

ORIGIN	Query Match	6.5%	Score	205.8	DB	28	Length	999
	Best Local Similarity	63.8%	Pred. No.	3.3e-39				
	Matches	400	Conservative	0	Mismatches	182	Indels	45
								Gaps
Qy	2090	CACAGAGTCTCCACATCCTCAGAACTTTTCAGTCTACTCAGGAACAGCAGCAGACATTTGC	2149					
Db	409	CAAAATTACCTGCTCCTTGATGACATTTGGAGTGTACCCAGAAACACAGACAAATTC	458					
Qy	2150	GGATGGAAGAAATGGTCACCTATTGCTGCAAGCTTCACCACCTATTTTCCATCATGATGATC	2209					
Db	469	ACAGGATGGAGGGTCACCAATTCATGATACCTCACCTATGTTTTCACATC	522					
Qy	2210	AGTATATGCTGAAGCACCACTGAACATTTGGGGCGTGAAGACGCAAGAACTAACGT	2269					
Db	523	AACATATTTGCTGAAGTGCCTGCTCGGAGCTGGAGGAATTAACGGGGAAAGAGTTTAATGT	582					
Qy	2270	GGGAGCAATTTTAAGSCACCTACAAGAAATTTCTCCAGCAGCAACATGTGGTGTCTCAATTTA	2329					
Db	583	GGGATTTCTTTCAGACAGCTTTAAGAAATTCACCAATCGTCATATGGTTTCAATTCG	642					
Qy	2330	GACCTGGTATCCAAAGCAGTTGACTTGCACCTTCTACTCATGTCTATGGGATCTTCAGCAAT	2389					
Db	643	GAAACAGGCTTCAAGAAGTTCAATTCATCTCCAGTCATGCTTATGGAGCTTCTAGTAAC	702					
Qy	2390	ATGCTATCTGCCAAACAGATTAATTCGSCCACTGGACCGCTATGCTGAAGAGCGGTTAAC	2449					
Db	703	ATGCAGTCCACAGCCAGTAATTTGGCGCTGAGA	736					
Qy	2450	AGGTCCATGCAAGAAATTTTCCAAGCACCAATAGCAACCATGGAAGCGAGTAAGTTATGTG	2509					
Db	737	-----TCAGCGGAGAGTGTTCAGGACCAACATCGACTATGGAGGCTGGTAGTTCGTATG	792					
Qy	2510	ATCGGAAATGCTGGCAAGTAGTCTTGATCTCTAAAGATCCATGCTCTGGCAGCGGATC	2569					
Db	793	ATCGAAGGATTCCTGCACAGTCAGGGCTCTATCCAAAAGAGCCCTATGCTGCAACACATC	852					
Qy	2570	TTCTGGAATGATGGATCCATCAACANTAGCAGCTTTCGCCACTATGGAACCTTCTTAGCA	2629					
Db	853	TTCTGAGATTTGATGGACTCATCAACAGCTCCAGACTTCACAAACTATCAAGAGCTTAACA	912					
Qy	2630	GGAAACAGATGGAGTCTCAACTTTCATA-----ATTCTCAGTAGTGA---CATAAATCAGT	2680					
Db	913	GGGCCAGATGGAACTCAACACAAACCTCTGGATTCAACAATATACACAGCATGATCAGT	972					
Qy	2681	ACAAAGGATCAACCGACATCATATG	2707					

973 ATATGGGTCTAGACATCATATG 999

Db

RESULT 8
BU974305

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BU974305 556 bp mRNA linear EST 22-OCT-2002
HB27110r BC Hordeum vulgare subsp. vulgare cDNA clone HB27110
5-PRIME, mRNA sequence.

BU974305
BU974305.1 GI:24225098

EST.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Hordeum.
1 (bases 1 to 556)
Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.
Barley ESTs from developing seeds
Unpublished (2002)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 556 Sd Error: 0.00
Plate: 27 row: 1 column: 10
Seq primer: M13rev.

FEATURES
source
Location/Qualifiers
1..556
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:239154"
/db_xref="taxon:112509"
/clone="HB27110"
/tissue_type="developing caryopsis"
/dev_stage="8-15 DAP (days after pollination)"
/lab_host="XL10-Gold"
/clone_lib="BC"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis, 8-15 DAP(days after pollination) Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

ORIGIN
Query Match 6.5%; Score 205.2; DB 13; Length 556;
Best Local Similarity 66.8%; Pred. No. 3; 7e-39;
Matches 354; Conservative 0; Mismatches 143; Indels 33; Gaps 3;

QY 1770 TGACGAGCTCTACCCAGAGAAAGAACCAAACTTGAAGTGAAGTCTGTGAAACAA 1829
Db 34 TGAGCGAAATTAACCTAGACAGAAATATACAAAGCAAGAGTGTGACAGAAACAA 93
QY 1830 GACCATGATGATGATCATCCCATGGATATGTTGACTGTACTTAAACACAGCATGA 1989
Db 94 TGGCATGACAGATGATTCCAATGGACATGTTGAAGTCTGTTGAGATGTGAGATAA 153
QY 1890 GAGCGAGCTTATGACTGAGACTGATGTTGTTCTGACATCAACCGTATTCATTCACAGCAAC 1949
Db 154 GAGACCACTGACAACTGAGACCAATTCGCTGATATAGTCATCAATCAAGATAAT 213
QY 1950 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2009
Db 214 GGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 273

2010 GTTTGACATTAATTTCCCAACAGAA---GTCCTTGGCATCCCAAGTACACAGAGAGATT 2066
Db 274 GTTTGACATTAATTTCCCAACAGAA---GTCCTTGGCATCCCAAGTACACAGAGAGATT 333
QY 2067 ACAGGGTCATTTGGCATTTGACCAACAAG-----AGTCTCCACA 2105
Db 334 ACAGATCGGTGAGACGCCACACACAGGTACTTATAGCATGCTCTCAATATTACAGAC 393
QY 2106 TCCTCAGAACTTTTCAGTCTACTCAGAAACAGCAGACACATTTTCGGATGGAAGAAATGCT 2165
Db 394 TCCTGGTTCATGTAATGCTTACACAGGAGCCACAGGCACATTCGGGCATGGGAGAAATTAGT 453
QY 2166 CACTATTGTCGAGCTCACCACCTATTTTTCATCATCATGATCATCATGATCATCATGATCAT 2225
Db 454 CACTGTTGCTGCAACTTCG-----CACAGACAGGATCAGCCTCTTGTCTGAGGC 504
QY 2226 ACCAACTGAACTTTGGGGCGGTAAAGACGCAACAAAGACTTAACGTTGGAGC 2275
Db 505 AACACCTGAGCGCTGGAGCCATGTGGGATCAAGAAGTGCATATGGAAC 554

RESULT 9
CG345663

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CG345663 787 bp DNA linear GSS 26-AUG-2003
OGMLW77TH ZM 0.7.1.5 KB Zea mays genomic clone ZMMBwa0614M09,
genomic survey sequence.

CG345663
CG345663.1 GI:34262929

GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 787)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends
Location/Qualifiers
1..787
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBwa0614M09"
/clone_lib="ZM 0.7.1.5 KB"
/note="Vector: pBESK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 6.5%; Score 205; DB 29; Length 787;
Best Local Similarity 61.5%; Pred. No. 4; 7e-39;
Matches 410; Conservative 0; Mismatches 230; Indels 27; Gaps 4;

QY 2455 CATGCAGAAATTTTCCACGACCAATAGCAACCATCGAAGCGAGTAGTATGTGATCGG 2514
Db 139 CAGCCGAGAAATGTTTCCACGACAGCGTTGGCTTATGAGCGCTGTGAGTGTACGATCAA 198
QY 2515 AGAATGCTGACAAAGTAGTCTGTATCTCTAAAGAAATCCATGCTCGCAGCGATCTTCTG 2574
Db 199 AGGATCCCTGACAGCCGCGCTTTTACCNAAGAGCCCATGCCTGCAACATCTCTCTG 258

ORIGIN	Query Match	Score	178.6;	DB 9;	Length	469;
	Best Local Similarity	65.3%;	Pred. No. 1.2e-32;			
	Matches	306;	Conservative	0;	Mismatches	139;
					Indels	24;
					Gaps	
Qy	1960	GATTGCTGTAATAGCTGCTCCAGGATGGTTCAGATTATCCATCAAGTGGTTTGACACT	2019			
Db	1	GATTGTACTATAATAGCTGCTCCGAGGATGGTCCAAATATATGTGTCAGTGGATTGACACC	60			
Qy	2020	AATTCCCAACAGAA---GTCTTTGGCATCCCAAAGTACACAGAAGGAGTTTACAGGGTTCAT	2076			
Db	61	ACTTCA CAGCAGAAGCGGCTTTTGGAAACCAGATAGTTACCAGAATGCAATGCCAGAATCGT	120			
Qy	2077	TTGGCTTAGCCACCAAG-----	2115			
Db	121	GTATCAGGCCACCCACACACGCTAGTCATACGCTCTCAATATACAGACTCTCGTGTTCAT	180			
Qy	2116	TTTCAGTCTTACTCAGGAACACAGACACATFTTCGGATGGAGAAATGGTCACTATTGCT	2175			
Db	181	GTAATGTCTATCCAGGAACCA CAGGCA CATTTCGGCATGGAGAAATTAGTCACTATTGCT	240			
Qy	2176	GCBAAGTCACCACTATTTTTCACATCATGTATGATCATATATGCTCAAGCAACCACTGAA	2235			
Db	241	GCBAAGTCGCGCACTATATAC CAGAC CAGGATTCAGCTCTTTCGCGAGCAACACCTGAG	300			
Qy	2236	CATTGGGGCGGTAAAGACGCAAAAGAGCTAACTGGGAGCAATTTTAAGGCCACTACAAGA	2295			
Db	301	CGCTGGAGCCATATATCGAGCAAAAGAGTCGATGGGGAACCTTTTCAAGGCACTTCCCGAG	360			
Qy	2296	AATTCTCCAGCAGCAATGTGTGTCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGG	2355			
Db	361	GATTATCAACCAACAGGTGTGTCTAGTTCGACCTTAGCAATTCGACGGTTGATTA	420			
Qy	2356	ACTTCTACTCATGTCATGGGAACTTCCAGCAATATATGCACTTCGCCAAC	2404			
Db	421	ACTTACTCATATGGCGGAGCTAATCGTATATATCCCACTCGCCAGC	469			

RESULT 11
 AC209720
 LOCUS
 DEFINITION
 491 bp DNA linear GSS 03-AUG-2003
 Oryza sativa (japonica cultivar-group) DNA, clone:NE8014_0.704_1A,
 3' flanking sequence of Tos17 insertion in rice strain NE8014,
 genomic survey sequence.
 ACCESSION
 AG209720
 VERSION
 AG209720.1 GI:32356910
 KEYWORDS
 GSS.
 SOURCE
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Sukayaka; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriarthoidae; Oryzaceae; Oryza.
 REFERENCE
 1 Miyao,A., Tanaka,K., Murata,K., Sawaki,H., Takeda,S., Abe,K.,
 Shinozuka,Y., Onosato,K. and Hirochika,H.
 Target Site Specificity of the Tos17 Retrotransposon Shows a
 Preference for Insertion within Genes and against Insertion in
 Retrotransposon-Rich Regions of the Genome
 Plant Cell 15 (8), 1771-1780 (2003)
 JOURNAL
 MEDLINE
 22779046
 PUBMED
 12897251
 REFERENCE
 2 (bases 1 to 491)
 Miyao,A., Onosato,K. and Hirochika,H.
 Direct Submission
 Submitted (16-APR-2002) Akio Miyao, National Institute of
 Agrobiological Sciences, Molecular Genetics; 2-1-2, Kannondai,
 Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp,
 URL:http://tos.nias.affrc.go.jp/, Tel.81-298-38-7020,
 Fax:81-298-38-7020)
 JOURNAL
 LOCATION/Qualifiers
 1. .491
 /orgasm="Oryza sativa (japonica cultivar-group)"
 FEATURES
 source

```

/mcl_type="genomic DNA"
/strain="NE8014"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="NE8014_0_704_1A"
/clone_lib="PCR product directly amplified from rice genomic DNA"
/notes="The 3' end of retrotransposon Tos17 was found immediately upstream of this sequence."
misc_feature 1
/note="This insertion point is base 139 in AF326768."

ORIGIN
Query Match 5.4%; Score 171.2; DB 29; Length 491;
Best Local Similarity 92.2%; Pred. No. 8e-31;
Matches 190; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

Qy 1 ATGGAGATGCTTCACGTA-GATCAGGAGGAGGCTGCTGTTGTTGGACGAACTGTATGCT 59
Db 103 ATGGAGATTGTTCCAGTAGATCAGAGGAGAGCTGCTGTTGTTGGACGAACTGTATGCT 162

Qy 60 TGCTCGTGTGGAACTGGTGCTGTACGGCAGTGTGTTGGAGCTGACAGGAGCGCTCGTCA 119
Db 163 TGCTCGTGTGGAACTGGTGCTGTACGGCAGNGTTGGAGCTGACAGGAGCACCTCGTCA 222

Qy 120 GGATGACGCGCTGGAAGCTGGTGATACGAAACGGGCACAAACCAATGCGAGCAATTTCTC 179
Db 223 GGATGACGCGCTGGAAGCTGGTGATACGAAACGGGCACAAACCAATGCGAGCAATTTCTC 282

Qy 180 CATAAGAGGGTATGTTGCTCTTCTTC 205
Db 283 CATAAGGTAATCATTTCTGTATTTTC 308

RESULT 12
ZZ411511

```

Z01-ANILION
genomic survey sequence.
BZ41I51l BZ41I51l GI:26045013 GSS.

Zea mays Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopshta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 641)

Whitelaw,C.A.; Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Sudiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics
Unpublished {2002}

Other_GSSs: OGACCAJTC
Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: Sheared ends.

Location/Qualifiers
1 .641
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWEMa0027112"
/clone_lib="ZW_0.7_1.5_KB"
/note="vector: pBCSK+ ; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES source

ORIGIN

```
/clone.lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      5.3%; Score 167.2; DB 28; Length 849;
Best Local Similarity 65.3%; Pred. No. 9.2e-30; Indels 21; Gaps 3;
Matches 298; Conservative 0; Mismatches 133;

QY 2720 ATGGAAGATTCATTCACATTCGACATTCGAGACATTCACGGCATCA---GCTGCATGATCTGC 2776
DB 741 ATGGAAGCCAAACCGCTGACACTGGAAGACTTGTCTCGGGTGTATTTCCAGCAAGACTTGC 682

QY 2777 ACAGACTTTACGCCACATCTTAGAGTTGGTGTGTCTTGGCTCTCTTGGCTGCGAGAGAA 2836
DB 681 GCAGGCTTTACGCCCTTACCCCTCGTGTGCGGTGTCTGCGGTTCATTGCTGCGAGGAGA 622

QY 2837 TTCCAAACTGCTCGGAGAACTGTGGCACAAATCTGTTTATAAGTTAGGAGTGTCAACAG 2896
DB 621 TCGCAACTGCTCTGGAACTGTGGGCCGCGAGCTGTGGTACAGACTAGGTGTTAAG 562

QY 2897 GATATCATCGCATCAGATGACAGAAAGAAACATTTTGAGGCCCTGAAATCTGGAAATG 2956
DB 561 GGACAAAGCCGCTGTATGCCAACAGAGACGAGAAACTACGAGACCTTGA-----514

QY 2957 TTTTACGAAATGGAATGCAATTCGATTCGATTTGGTTTCTGTAGCT-----CCAGTGCAGATT 3010
DB 513 GCTGGCAGATGGAACACCTTGCAGTTGGTTTCTGTAGCTCTGCTGCAATCTGGAGT 454

QY 3011 TTTTATCAGCGAGAAACAGCATGATGCTAAATCTTTGGACGAGCAAGGGTAAATGGTTTC 3070
DB 453 ACCGGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394

QY 3071 ATCCCTTGGATGGTTTGTGACAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3130
DB 393 ATCCGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 334

QY 3131 TTACTACATCAGTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3162
DB 333 TCAGTGAATAGTGACAGCAACGAGTACATG 302

RESULT 14
CC603916/c
LOCUS
DEFINITION
  OGWM55TV ZM 0.7 1.5 KB Zea mays genomic clone ZM80296U13,
  genomic survey sequence.
ACCESSION
  CC603916
VERSION
  CC603916.1 GI:31965337
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
    Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
    Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
    Consortium for Maize Genomics
    Unpublished (2002)
    Other_GSSs: OGWM55TH
    Contact: Cathy Whitelaw
    TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TP
    Class: sheared ends.
    Location/Qualifiers
      1..766
        /organism="Zea mays"

FEATURES
  source
```

```
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM80296U13"
/clone.lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
```

```
ORIGIN
Query Match      4.8%; Score 152; DB 29; Length 766;
Best Local Similarity 57.3%; Pred. No. 4.8e-26;
Matches 434; Conservative 0; Mismatches 255; Indels 69; Gaps 6;

QY 2094 AGAGTCTCCACATCTCTCAGAACTTTTCAGTCTACTCAGAAACAGACAGACATTTGCGGAT 2153
DB 698 AAAATTACTTCCCTTGCATCTACTCTGAGTGTCCCGAGAACCAACGACACGATTCACAG 639

QY 2154 GGAAGAAATGTTCACTATTGTTGCTGCAAGCTACCACTATTTTTCATCATCATGATGATG 2213
DB 638 GGCAGATGAAGTCCACATCGCATGTACCTGCGCTATATTTTCAAACT-----ATCAACA 585

QY 2214 TATTCTGAGACACCAATCTGAACATTTGGGCGGTGAAGCGCAAGAAAGCTAACTGGGA 2273
DB 584 TATTCTGAAAGCGCTGCTCGGAGCTGGAGGAGCAGCGGAAAGAGCTAATGTGGGA 525

QY 2274 GCAATTTAAGGCCACTTACAAGAAAT-----TCTCCAGCAGCAACATGTGCTGCTCAATT 2327
DB 524 TTCTTTCAAGCAGCTTCAAGGAATCCCACTACCAACCGCTCATATGTTTCAATT 465

QY 2328 TAGACTGGTATCCAAAGAGTTGACTTCTACTCTACTCATGTCTATGCTGAGGATCTTCCAGCA 2387
DB 464 CGGAACACAGGCTTCAAGAAAGTTGATTTCAGCTCCCAATCATG-----C 423

QY 2388 TTATGCTATCTCCCAACACAGTAATTTGGCCACTGACCGCTATGCTGAAAGACGGTTAA 2447
DB 422 ATATGAGCTTACACGCCAGTAATTTGTGGCGTAGAT-----386

QY 2448 CCAGGTCCATGCAAGAAATTTTCCAAAGCACAATAGCAACCAATGAGAGCGAGTAAGTTATG 2507
DB 385 -----CAGCCGAGAAGTTTCCAAAGCAGCAGCTTGGCTATGAGGCTGGTAGGCTGTA 333

QY 2508 TGATCGGGAATGCTGGACAGTAGTCTTGTATCTTAAAGAAATCCATGCTCCGACGCA 2567
DB 332 CGATCAAGAGGATCCCTGGACAGCCAGGCTTTTCCCAAGAGAGCCATGCTTGCACACA 273

QY 2568 TTTTCTGAGAATGATGGATCCATCAACATTAGCAAGTTCCTCCAACTATGGAACCTTTTAG 2627
DB 272 TCTCTGAGGCTGATGGACTCATCAACAGCTCCAGGCTTCAAACTACCAAGAGACTAA 213

QY 2628 CAGGAACGAGTGGAGTCTCAACTTCAATTTCTAGTATGCACATATCATCATGACATGAG 2687
DB 212 TAGGCGGAGATGAGGCTTCAAAACGCAAGCTCTCGGCTCGCAGTACACACGATGAGCA 153

QY 2688 ATCAACCCAGCACAATCATATGCGAGTAACCTGAATGGAAGATTTCCATTGACATTCGAAGA 2747
DB 152 TT-----ACAATGCGTCCACCGGACCGTCAATATGGAAGCCAAACCGCTGACACTGGAAGA 99

QY 2748 CTTATCAGGCAATCA---GCTGCTGATGATGACAGACCTTTACGCCCACTCTTAGAT 2804
DB 98 CTTGCTCTCGGCTGATTTTCCAGCAAGACTTGGCGAGGCGCTTTACGCCCTCACCCCTGCT 39

QY 2805 TGTTGTGCTTGGCTCCTCTGCTGCGAAGGAATTTGCAA 2842
DB 38 CGGTGCTCGGTTTCATCTGTCGACGACGAGATGCGAA 1
```

```
RESULT 15
BH880296/c
LOCUS
DEFINITION
  BH880296
  ht52dl1.b1 WGS-ZnaysF (JM107 adapted methyl filtered) Zea mays
  genomic clone ht52dl1 5', genomic survey sequence.
ACCESSION
  BH880296
VERSION
  BH880296.1 GI:22116193
```

KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Bailly, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)

TITLE
JOURNAL
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ht52 row: d column: 11
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 635.

FEATURES
source
1..635
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ht52d11"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
/notes="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(x/y reads in M13mp19, b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

ORIGIN
Query Match 4.4%; Score 139.4; DB 28; Length 635;
Best Local Similarity 65.6%; Pred. No. 5.6e-23;
Matches 254; Conservative 0; Mismatches 121; Indels 12; Gaps 3;

QY 1684 CAGATGGAACCGAAACTCTGTCTGAGTCACCTCGCAAGGTTTCTCCAGCTGAGCAT 1743
Db 383 CTGCTTGTATACAAATATTGTACCCCTCATCTCGTTACCAAGGTTTCTCTTGTAAACAA 324
QY 1744 GATATCCAAATTATGTCGACCTTCATGACGAGCTACCCCAAGAGAAAGAGACAA 1803
Db 323 AAGATCCAGATGTGTCTGCTCCACAGAGAATATACCAAGATAAGA-----AA 270
QY 1804 AAATTTGAAGTACTCTGTGAAACACAGCCATGATAGATGACATCCCATGGATTTGTT 1863
Db 269 GGAAGCAGAAGGTGCAACGAGAACACAGATGTATAGACACTTCCCCATGGACATTGTT 210
QY 1864 GAATGCTAGCTAAACACAGCATGAGGACGCTTATGACTGAGACTGATTTGTTCTGAC 1923
Db 209 GAATCTTCCAGAAATCAGCAGAGACAGTGTATGACTCAGACCAATCTTTGGAA 150
QY 1924 ATCAACCGTATTCAATCCAGACAACTGTGTGATGATTTGTATATAGTAGTGCACAG 1983
Db 149 AACTATCATACTCAACCCAGGTA--GCTCAAGTTGATTTGCTGCAATTTGCAGCCAG 93
QY 1984 GATGGTTCAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2043
Db 92 GGTGTGTCCTCCCAATGATCAATGATGATGATGATGATGATGATGATGATGATGATG 9
QY 2044 TCCCAAGATACACAGAGGAGTTACAG 2070
Db 35 TCGGAAGTAAGCAGAGTCCCTTACAG 9

Search completed: May 17, 2004, 01:43:09
Job time : 7754 secs

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 17, 2004, 01:47:00 ; Search time 10018 Seconds
(without alignments)
4573.127 Million cell updates/sec

Title: US-09-721-114-2
Perfect score: 5526
Sequence: 1 MEIVAVDQSGARVVGTCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgr2 1/USPTO_spool/US09721114/runat 14052004 092251 4115/app_query_faasta_1.1223
-DB=GenEmbl -FAST=fastap -SUFFIX=rsge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09721114 @CGN 1 1 5626 @runat 14052004 092251 4115 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WASN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_ste.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5526	100.0	4310	6	BD083699	BD083699 Novel gen
2	5516	99.8	3896	8	AF326768	AF326768 Oryza sat
3	5191	93.9	9455	6	BD083700	BD083700 Novel gen
4	5191	93.9	150594	8	AP001859	AP001859 Oryza sat
5	5179	93.7	3626	8	AK103364	AK103364 Oryza sat
6	393	7.1	3291	6	AX590041	AX590041 Sequence
7	298	5.4	8648	6	AR368885	AR368885 Sequence
8	298	5.4	8648	8	AF319968	AF319968 Arabidops
9	298	5.4	17341	6	AR368884	AR368884 Arabidops
10	298	5.4	101665	8	ATFL5N18	ATFL5N18 Arabidops
11	241.5	4.4	7036	10	MUSCYCLOA	L04289 Mus musculu
12	234.5	4.2	32367	10	AF520421S2	AF520422 Mus muscu
13	234.5	4.2	22782	2	EX842684	EX842684 Danio rer
14	233	4.2	184271	3	AC009256	AC009256 Drosophil
15	233	4.2	198961	3	AC010919	AC010919 Drosophil
16	233	4.2	323844	3	AE003501	AE003501 Drosophil
17	227.5	4.1	212955	2	AC136676	AC136676 Rattus no
18	225	4.1	160827	3	AC099025	AC099025 Drosophil
19	223.5	4.0	193262	3	AC007579	AC007579 Drosophil
20	223.5	4.0	260367	3	AE003808	AE003808 Drosophil
21	223	4.0	3114	9	HSATRX10	U72909 Homo sapien
22	223	4.0	4808	3	AY075323	AY075323 Drosophil
23	223	4.0	7479	9	AB102641	AB102641 Homo sapi
24	223	4.0	10216	9	HSU72938	U72938 Homo sapien
25	223	4.0	10330	6	AX828379	AX828379 Sequence
26	223	4.0	10330	9	HSU72937	U72937 Homo sapien
27	223	4.0	10452	6	AX336451	AX336451 Sequence
28	223	4.0	10452	6	AX828404	AX828404 Sequence
29	223	4.0	10452	9	HSU72936	U72936 Homo sapien
30	223	4.0	47323	9	AC005937	AC005937 Homo sapi
31	223	4.0	84474	9	AL662854	AL662854 Human DNA
32	223	4.0	103780	9	AL773541	AL773541 Human DNA
33	223	4.0	110000	2	BX276116_03	Continuation (4 of
34	223	4.0	135982	9	HS34606	Z84487 Human DNA s
35	223	4.0	153440	3	AC105263	AC105263 Drosophil
36	223	4.0	156272	9	AL669830	AL669830 Human DNA
37	223	4.0	178688	2	AL713893	AL713893 Homo sapi
38	223	4.0	276131	3	AE003543	AE003543 Drosophil
39	222	4.0	10652	10	AF441786	AF441786 Mus muscu
40	221	4.0	3829	9	HSXNP08	AF000160 Homo sapi
41	221	4.0	6542	9	AB014524	AB014524 Homo sapi
42	221	4.0	6655	9	AY099469	AY099469 Homo sapi
43	221	4.0	10358	9	HSU75653	U75653 Human zinc
44	221	4.0	151865	2	AC025197	AC025197 Homo sapi
45	221	4.0	347050	3	PFA929351	AL929351 Plasmodiu

ALIGNMENTS

RESULT 1

BD083699
LOCUS
DEFINITION Novel gene participating in response to brassinosteroid.
ACCESSION BD083699
VERSION BD083699.1 GI:22629309
KEYWORDS JP 2001327287-A/1.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Hirochika,H., Yamazaki,T. and Miyao,A.
TITLE Novel gene participating in response to brassinosteroid
JOURNAL Patent: JP 2001327287-A 1 27-NOV-2001;
NATL INST OF AGRONOMICAL RESOURCES, BIO ORIENTED TECHNOLOGY RESEARCH ADVN JAPAN FORAGE SEED ASSOCIATION
COMMENT OS Oryza sativa (rice)
PN JP 2001327287-A/1
PD 27-NOV-2001
PF 19-MAY-2000 JP 2000149106
PI HIKOHiko HIROCHIKA,TOKIO YAMAZAKI,AKIO MIYAO
PC C12N15/09,C12N15/00
CC Novel gene participating in response to brassinosteroid FH
Key Location/Qualifiers
FT CDS (655)..(3828).
FEATURES
source
1..4310
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4310
Score: 5526.00 Matches: 1057
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-721-114-2 (1-1057) x BD083699 (1-4310)
Qy 1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
Db 655 ATGCGAGATTGTTGTCAGTAGATCATGAGGAGCTCGTGTTCGTGGGACGAACCTGTATGCTT 714
Qy 21 AlaArgGlyGlyThrGlyAlaValAlaProValIleuGluIleuThrAlaThrProArgGln 40
Db 715 GCTGTGTGTGGAACTGTGTCTGTAGCCGCTGTGGAGCTGACGACGAGCTCTGCTCAG 774
Qy 41 AspAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
Db 775 GATGCGAGCGCTGAAGCTGGGTGTAGACGAACGGCACACACCAATGCGAGCATTTCTCC 834
Qy 61 IleArgGlyTyrValAlaIleuLeuGlnLysLysAspProIlePheCysSerLeuSerArg 80
Db 835 ATAAGAGGGTATGTGCTCTCTTCAGAGAGGATCCAAAATCTGCTCTCTATCTCGG 894
Qy 81 IlePheHisAspGlnLysLysCysAspGluHisIleValAsaSerSerPheSerVal 100
Db 895 ATTTTCATGACCCAGAAAAAATGTGTATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954
Qy 101 AlaIlePheArgTrpIleAspCysSerLysCysIleuAspLysLeuLysThrSerAspAsn 120
Db 955 GCAAGATTTCGACGATGGGATGCTCTGCAAGTGTTCGATAAGTTGAAAACCTCAGATAAT 1014
Qy 121 GlyThrAlaProArgThrIleuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db 1015 GGAACAGCACCAAGAACTCTTCCCGCCAAAGCAGATGGCAGATGATGGTGTCTCATC 1074
Qy 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160

1075	Db	ACATTTGTTCCGAGACACTTTTGTGGCTCTGCTAGTGTGGTTTCCAAAAAGTGCTCTCTAGC	1133
161	Qy	ThrGlnSerSerGlnGlyLysAsnAlaAaspArgSerThrLeuProLysSerValGlnGlu	180
1135	Db	ACACAATCATCTCAAGCGGAAGATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA	1194
181	Qy	GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr	200
1195	Db	GGCAATGACATCCAAATGCAATGCCTCTTGGCAGAAATGGAGCTGCTGAGGCCAATACT	1254
201	Qy	AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyAspValAlaAlaAsnVal	220
1255	Db	GATTCCACCAATGAAAGATTGTCAGGCGCCAGGCCCAAAATTTATGATGGCAGCAAAATGTC	1314
221	Qy	SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp	240
1315	Db	TCTGAGGACACACTTCTGTGTGATGTGGGGCTTTTACCTGAAGTTCCCCAGATTACATGG	1374
241	Qy	HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal	260
1375	Db	CACATGAGAGTAATGGTGCAGATCACTCCATCCACTCCAAAACCTTTCTGAAGTGGTC	1434
261	Qy	LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys	280
1435	Db	CTCRAAGAAATCAAGATGAAATGGAAATCTGAAGACACTCTTGTTGCTGAGCAGTGC	1494
281	Qy	AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln	300
1495	Db	AAATTTGACCAAGATCTTAACCCAAATGTTCTGAAAGGAACTGTATCAGGTGCTGAGACAG	1554
301	Qy	CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn	320
1555	Db	TGCAATTTTGACCAAGATCCGAACACAGTGTCTGGCGCAAAATGTGAGCAGATCTGCAAT	1614
321	Qy	GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys	340
1615	Db	GAGCCATGTGAAGAGTTGTTCTCAAAAGAGCTCCAAATCTAAGAGGAAGACGGATAG	1674
341	Qy	LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp	360
1675	Db	AACTTGATGAAGAAGCAGCAGCACAGCAAGAAACGACCTGCCCGGCTGATGTTTCCAGAT	1734
361	Qy	AlaLysLeuCysArgArgLysProLysLysValArgLeuSerGluIleIleAsnAla	380
1735	Db	GCAAAGCTTTGTCCGAGAAGCCAAAAAAGTGCGGCTTCTATCAGAAATTTAAATGCT	1794
381	Qy	AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys	400
1795	Db	AACCAAGTTGAGATTTTAGAGTGCAGAGTTTCATCGTGAAAATGCCCGCTGATCCCTGT	1854
401	Qy	GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer	420
1855	Db	GAGGATCATAGTAGTACCATCCGGTCCCGATGGAAGTAAAGCATGGATATTCCTGTAGC	1914
421	Qy	AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysIyr	440
1915	Db	AACCATACAGTGGAGAGATGGGTTTAAATCAAGTAGAACAAGACAAAACGCCAANTAC	1974
441	Qy	SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg	460
1975	Db	TCTGATGTTGTAGATGATGGATCATCACATTATGAACCTGGCTGAATGGAAAAAAGAAAAGA	2034
461	Qy	ThrGlySerValHisIleThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal	480
2035	Db	ACTGGAAGTGTGCATCACACAGTTGCTCTATCCAGCTGGGAATTTGAGCAACAAAAAGTG	2094
481	Qy	ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr	500
2095	Db	ACACCCACTCGGAGTACTCAGCATGATGATGAGAAATGATCTGAAATGGTCTTGACACA	2154
501	Qy	AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer	520
2155	Db	AATATCATCAAGACAGATCTGTGACGATGTATCAGAATCTCCACAACATGGTGCTCA	2214

QY 521 SerIysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
DB 2215 TCAAGGGGAACACAGGGGTTTGTAGTAAGGGGAACACATTCAGCTGTAGTACCA 2274
QY 541 TyrGlyGlyGlySerThrArgAsnGlyGlyAsnIleHisValLeuSerAlaGluAspGln 560
DB 2275 TATGGTGGTGAAGACACAGAAATGGTCAAGACATATCATGTACTCAGCGCAGAGATCAA 2334
QY 561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
DB 2335 TGCAGATGGAAACCGAANAATCTGTCTGAGTCATCTGGCAAGGTTTCTCAGCTGAG 2394
QY 581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLys 600
DB 2395 CATGATATCCAAATTTATGTCTGACCTTCATGACAGAGTCTACCCAGGAAGAAAAGAAG 2454
QY 601 GlnLysLeuGluValThrArgGlyLysGlnThrMetIleAspAspIleProMetAspIle 620
DB 2455 CAAPAACTTGAAGTACTCTGTGAAGAAACACCATGATGATGATGATGATGATGATGAT 2514
QY 621 ValGluLeuLeuAlaLysAsnGlnHisGlyArgGlnLeuMetThrGluThrAspCysSer 640
DB 2515 GTTGAATCTGTAGTAAACACAGCATGAGAGGAGCTTATGATGATGATGATGATGAT 2574
QY 641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
DB 2575 GACATCAACCGTATTCAATCCAAAGACACATGCTGATGATGATGATGATGATGATGAT 2634
QY 661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
DB 2635 AAGGATGTTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694
QY 681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
DB 2695 GCATCCCAAGTACACAGAGGAGTACAGGGTCAATTTGGCATTGACCAACAGAGTCT 2754
QY 701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuArgMetGluGlu 720
DB 2755 CCACATCTCAGAACCTTTCAGTCTACTCAGAAACAGCAGACATTTGCCATGGAGAA 2814
QY 721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
DB 2815 ATGGTCACTATTGCTGCAAGCTCACCACCTATTTTCATCATCATCATCATCATCAT 2874
QY 741 GluAlaProThrGluHisTyrGlyArgLysAspAlaLysIleLysLeuThrTyrGluGlnPhe 760
DB 2875 GAAGCAACCACTGAACATTTGGGCGCGTAAGGACGCAAGAGCTAACGTTGGGAGCAATT 2934
QY 761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyValAlaGlnPheArgProGlyIle 780
DB 2935 AAGCCCACTACAGAAATTTCTCAGCAGCAACATGTTGTTCTCAATTTAGACCTGGTATC 2994
QY 781 GluAlaValAspLeuThrSerThrHisValMetGlySerSerSerSerSerSerSerSer 800
DB 2995 CAAGCAGTTGACTTGACTTCTTCTACTGATGATGATGATGATGATGATGATGATGAT 3054
QY 801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
DB 3055 CAACCAAGTAAATGGGCACTGGACCGCTATGCTGAAAGACGCGTTAACCGGTTCCATGCA 3114
QY 821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
DB 3115 AGAAATTTTCCAGACCAATAGCAACCATCGAAGCGAGTAAGTTATGTATGTCATGCGA 3174
QY 841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
DB 3175 GTGGACCAAGTAGTCTGTATCTCTTAAGAAATCATGCTCCGACGATCTCTCTGAAAGT 3234
QY 861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
DB 3235 ATGGATCCATCAACATTTAGCAAGCTTCCCACTATGAACTTCTAGCAGAACCAAGATG 3294

QY 881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
DB 3295 GAGTCTCAACTTTCATAATTTCTCAGTATGCACATAATTCAGTACAAAGGATCAACACACA 3354
QY 901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
DB 3355 TCATATGGCAGTAACCTGAATGGAAGATTCATTCATGACATTCGAGACTTATCACGGCAT 3414
QY 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
DB 3415 CAGCTGCATGATCTGCACAGACCTTTACGCCACATCTTAGAGTTGGTGTGCTGCTCC 3474
QY 941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
DB 3475 TTGCTGCAGAGAGGAATTCGAAACTGGTGGAGAACTGTGGCAGACAACTCTGTTATAAG 3534
QY 961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
DB 3535 TTAGGNGTGTCAACAGGAATACATCGCATCAGATGACAGAAAGAAACATTTTGAAGCC 3594
QY 981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
DB 3595 CTGAATTTCTGGAATGTTTTCAGCAAAATGGAATGCAATTCGAGTTGGTGTGCTGCTCC 3654
QY 1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
DB 3655 AGTCGAGATTTTTCATGCGAGGAGACAGCATGCTCAATCTTGACAGAGCAAGGCT 3714
QY 1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
DB 3715 AAAATGGTTTCATCCCTTGGATCGTTCGTCAGACAGCATATCTGTATTAACAAAGAAC 3774
QY 1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
DB 3775 CCAGCTGATTTTACTACATTCAGTAACAGTAACAGATATATGATATACCGC 3825

RESULT 2
AF326768 3896 bp mRNA linear PLN 05-SEP-2001
LOCUS Oryza sativa embryonic flower 1-like protein mRNA, complete cds.
DEFINITION AF326768
ACCESSION AF326768
VERSION AF326768.1 GI:15430698
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
REFERENCE 1 (bases 1 to 3896)
AUTHORS Aubert,D., Chen,L., Moon,Y.H., Martin,D., Castle,L.A., Yang,C.H.
and Sung,Z.R.
EMPI, a novel protein involved in the control of shoot architecture
and flowering in Arabidopsis
JOURNAL Plant Cell 13 (8), 1865-1875 (2001)
MEDLINE 21380417
PUBMED 11487698
REFERENCE 2 (bases 1 to 3896)
AUTHORS Moon,Y.-H., Chen,L. and Sung,Z.R.
Direct Submission
TITLE Submitted (06-DEC-2000) Plant and Microbial Biol. Dept., University
of California, Berkeley 361 Koshland Hall, Berkeley, CA 94720, USA
FEATURES
source
1..3896
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
241..3414
/note="EMPI"
/codon_start=1
/product="embryonic flower 1-like protein"
/protein_id="AAK98529.1"
/db_xref="GI:15430698"
/translation="MEIVAVDQEGARVVGTNCMLARGGTGAVPVLTLTATPRQDAAA"

PAGVDEPAHQCEHFISIRGVALLQXKDPKFCSLISRI1PHDQCKCDEHKASSPSPVAC
FRMDCSKLIDKSTSPGQVNDAPRLKPAQNGKSGTSDGCSIT1VRSTF7VSGVQSKVPS
TQSCQKXNDR1KSTSPGQVNDKCNKSPGNGKGAENTDTPSKD1QJQAPNDVAA
NBSZDNT1VSGALPVPQ1TWH1EVNGADPFPSTPKLSEV1KXNDENGKTEETLY
AEQCN1TKDPNPMSPKQEVQACQNL1KDPKQPS7QCEQ1CNECPSEV1VLRKSKS
KCT1KMLKQKQFSKSKRTAQADVSDAK1LKRKPKKVALLSGI1VNAQVQVSDSEVH
RENAAD1CEDDRST1PVPMSVD1PVSXN1TGVGBDQSKXNKTQKSYDWDGGSGL
MLNANGKKT1GSGVTH1VAPAGNLSNKNKVT1PASTQ1HDBNDTNG1LHMTKH1DVC
QHYVS1STORCSKGT1AGLSKSHSAAT1VSGTES1TNDN1HLSABEQCMQMBE
NSVLHSKASVPSBAHD1Q1MSDLHQSLQKPKKQK1VTRKCTK1DID1PMD1VELL
AKQHQERQ1MATED1CSD1NRI1QSKT1ADDDQ1V1AAKQSGV1ASVPTD1NSQKLS
QSTQ1QKQHLAULT1QSP1NFTQ3QV1QKQTHLMKXGSDV1IASSP1TSHDDQ1YIA
EAPTEHVRKDKAL1TWEQ1KAT1RNSPAPT1CQAQ1REG1QAOVLTS1THVWSSGNTA
SROBP1APLDR1YAER1VANO1HARN1PST1AT1MKE1KCDRNNQAV1V1PKSPN1PATH
LLRMD1DSTLSP1FNT7P1RNVQMSQLN1QKQ1YANN1KGTST1SYSGNLMK1PIUT
EDLS1RQJLHLD1PLRPH1F7R1VQSG1LQK1LHANN1SNCQ1QSGY1KGTGT1SHQM
NRKHE1FALANG1MFSF1AKNAL1QV1AL1S1VSGV1SAY1FALNS1QAQST1RKGKMW1PLDRFV
ROD1CI1TKNKPAD1T1S1NDN1EMDYR1"

ORIGIN

Alignment Scores:

Pred. No.:

FILED NO.:
score: 5516-00

```
Score: 3516.00
Percent Similarity: 99.91%
Matches: 3516.00
Conserved: 3516.00
```

Percent Similarity: 99.91%
Best Local Similarity: 99.81%
Mismatchatches: 0
Conservative: 0

BEST LOCAL SIMILARITY: 99.81%
 MISMATCHES: 0
 INDELS: 0
 QUERY MATCH: 99.82%

Query Match:
DB:

h: 3896

1055

Derivative: 1

atches: 1

10

00

US-09-721-114-2 (1-1057) X AF326768 (1-3896)

Qy	1	MetGluIleValAlaValAspClnGlnClyAlaArgValValGlyThrAsnCysMetLeu	20
Db	241	ATGGAGATTGTTGCAGTAGATCAGAGGGAGCTCGTGTGTGGGACGAACGTGTATGCTT	300
Qy	21	AlaArgGlyGlyThrGlyValaValAlaProValLeuGluLeuThrAlaThrProArgGln	40
Db	301	GCTCGTGGTGAACCTGGTCTGTAGCGCCAGTGTGGAGCTGACAGCGACGCTCGTCAG	360
Qy	41	AspAlaAlaAlaGluAlaGlyValAspClnProAlaGlnHisGlnCysGluHisPheSer	60
Db	361	GATGCAAGCCGCTGAAGCTGGTGATAGCAACCGGCAACACCAATGCGAGCAATTCCTCC	420
Qy	61	IleArgGlyTyrValAlaLeuLeuGlnIlysisAspProLysPheCysSerLeuSerArg	80
Db	421	ATAAGAGGGTATGTTGCTCTCTTCAGAAAGAAGCATCCAAAAATTCGTCTCTATCTCG	480
Qy	81	IlePheHisAspGlnIlysisCysAspGluHisIlyValaserSerSerProPheSerVal	100
Db	481	ATTTTCCATGACAGAAAATAATGTGATGACACAAAGCTAGTTCAGGCCATTTCTGTAT	540
Qy	101	AlaLysPheArgArgThrAspCysSerIysCysLeuAspIlysisLeuIlyThrSerAspAsn	120
Db	541	GCAAGATTTTCGACGATGGGATTGCTCGAAGTGCCTTGGATAGTTGAATAACTTCAGATAT	600
Qy	121	GlyThrAlaProArgThrLeuProAlaIysGlnAsnGlyThrSerAspGlyCysSerIle	140
Db	601	GGAAACAGCACCAAGAACCTCTTCCCGCAAGCAGAAATGGCACAGTGTGTGCTCCATC	660
Qy	141	ThrPheValArgSerThrPheValProIleValSerValGlySerGlnIlysisValSerProSer	160
Db	661	ACATTTGTTTCGAGGACATTTTGTGCTGCTAGTGTGGTTCCTCCAAAAGTGTCTCTCAGC	720
Qy	161	ThrGlnSerSerGlnGlyIysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu	180
Db	721	ACACAATCATCTCAAGGGGAAGAATGCTCATAGATCAACTCTTCCAAAAGAGTGTGCAAGAA	780
Qy	181	GlyAsnAspSerIysCysAsnAlaProSerGlyIysAsnGlyValAlaGluAlaAsnThr	200
Db	781	GGCAATGATCTCCAAATGCAATGCGCTCTCGCAAGAATGAGAGTGTGTGAGGCAATACT	840
Qy	201	AspSerProMetLysAspLeuGlnGlyProAlaGlnIleNtyrAspValAlaAlaAsnVal	220
Db	841	GATTCACCAATGAAGATTTCGAAGGGCAGCCCAAAATATGATGTGGCAGCAAAATGTC	900

QY 581 HisAspIleGlnIleMetSerAspLeuHisGlnSerLeuProTyrLysLysLysLys 600
DB 1981 CATGATATCCAAATATGTCGACCTTCATGAGCAGAGTCTACCCAGAGAAAGAG 2040
QY 601 GlnLysLeuGlnValThrArgGluLysGlnThrMetIleAspIleProMetAspIle 620
DB 2041 CAAAACTTCAGAGTCTCGTGAAGAAACAGACCATGATGATGATCCCTCGGATAT 2100
QY 621 ValGlnLeuLeuAlaLysAsnGlnHisGlnArgGlnLeuMetThrGlnThrAspCysSer 640
DB 2101 GTTGAACTGTAGCTAAACACAGCATGAGAGCAGCTTATGACTGAGAGTATGTTCT 2160
QY 641 AspIleAsnArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAla 660
DB 2161 GACATCAACGATATTCATCAAGACACACTGCTGATGATGATGTTGTAATAGTCTGC 2220
QY 661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
DB 2221 AAGGATGGTTCAGATATGATCATCAGTGTGTTGACACTAATTCACCAACAGAGTCTTG 2280
QY 681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
DB 2281 GCATCCCAAGTACACAGAGGAGTTACAGGGTCAATTTGGCATTTGACCAACAGAGTCT 2340
QY 701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuArgMetGluGlu 720
DB 2341 CCATCCTCAGAACTTTCAGTCTACTCAGCAACAGCAGACATTTGCCGATGGAGAA 2400
QY 721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspGlnTyrIleAla 740
DB 2401 ATGGTCACTATTCGTGCAAGTCCACCACTATTTTCATCATCATGATCATCATGATTC 2460
QY 741 GluAlaProThrGluHisIleTyrGlyValGlyAspAlaLysLysLeuThrTyrGlnPhe 760
DB 2461 GAAGCACCACTGAACATTGGGCGGTAAGAGCGCAAGAGCTTAAGTGGGAGCAATTT 2520
QY 761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
DB 2521 AAGGCCACTACAGAAATTCCTCAGCAGCAACATGTTGCTGCTCAATTTAGACCTGATC 2580
QY 781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
DB 2581 CAAGCACTTGACTGACTTCTACTCATGTCATGGGATCTTCAGCAATTTATGATCTGCG 2640
QY 801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
DB 2641 CAACCAAGTANTGGCCACTGGACCGTATGCTGAAGAGCGGTTAACCAAGTCCATGCA 2700
QY 821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgAsn 840
DB 2701 AGAAATTTCCAGCACAAATAGCAACCATGSAAGCGAGTAAGTTATGTGATCGAGAAAT 2760
QY 841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
DB 2761 GTGGACAGTAGTCTTGTATCTCTAAAGAAATCCATGCTCGCAGCGCATCTCTGAGATG 2820
QY 861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
DB 2821 ATGGATCCATCAACATTAGCAAGCTTCCCAACTATGGAACCTCTAGCAAGAACAGATG 2880
QY 881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
DB 2881 GAGTCTCAACTTCAATAATCTCAGTAGTCACATAATCAGTACAAAGGATCAACCAACACA 2940
QY 901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
DB 2941 TCATATGGCAGTAACCTGAATGGAAAGATTCATTCATTCGAGACTTATCAAGCAT 3000
QY 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
DB 3001 CAGCTGCATCATCTGCACAGACCTTTACGCCACATCTCAGAGTTGCTGCTGCTGCC 3060
QY 941 LeuLeuGlnLysGluIleAlaAsnTyrSerGluAsnCysGlyThrGlnSerGlyTyrLys 960

DB 3061 TTGTCGACAGAGGAATTCGAAACTGGTCGAGAACTGGGCACACAAATCTGTTTATAAG 3120
QY 961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
DB 3121 TTAGAGTGTTCACAGGAATACATCGCATCAGATGAACAGAAAGCAATTTTGAGGCC 3180
QY 981 LeuAsnSerGlyMetPheSerAlaLysTyrAsnAlaLeuGlnLeuGlySerValSerSer 1000
DB 3181 CTGAATTCGGAATGTTTTTCAGCAAAATGGAATGCAATTCAGTTGGGTTCTGTATGCTCC 3240
QY 1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTyrThrArgGlyLysGly 1020
DB 3241 AGTGCAATTTTTTATCAGCAGGAACAGCATAGCTCAATCTTGGACCAAGCAAGGGT 3300
QY 1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
DB 3301 AAAATGGTTTCATCCCTTGCATCGTTTGTGAGACAGATATCTGTATAACTAACAGAAC 3360
QY 1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
DB 3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGATTTACCGC 3411
RESULT 3
BD083700 9455 bp DNA linear PAT 27-AUG-2002
LOCUS Novel gene participating in response to brassinosteroid.
DEFINITION BD083700
ACCESSION BD083700.1 GI:22629310
VERSION JP 2001327287-A/2.
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 9455)
AUTHORS Hirochika,H., Yamazaki,T. and Miyao,A.
TITLE Novel gene participating in response to brassinosteroid
JOURNAL Patent: JP 2001327287-A 2 27-NOV-2001;
NATL INST OF AGRICULTURAL RESOURCES, BIO ORIENTED TECHNOLOGY
RESEARCH ADVANCEMENT INSTITUTION
COMMENT OS Oryza sativa (rice)
PN JP 2001327287-A/2
PD 27-NOV-2001
PF 19-MAY-2000 JP 2000149106
PI HIROHIKO HIROCHIKA, TOKIO YAMAZAKI, AKIO MIYAO
PC C12N15/09, C12N15/00
CC Novel gene participating in response to brassinosteroid FH
Key source Location/Qualifiers
FT source 1..9455
FT Location/Qualifiers
1..9455
/organism='Oryza sativa (rice)'.
/mol_type='genomic DNA'
/db_xref='taxon:4530'
FEATURES
Source
ORIGIN
Alignment Scores:
Pred. No.: 1,06e-302 Length: 9455
Score: 5191.00 Matches: 1055
Percent Similarity: 68.06% Conservatives: 0
Best Local Similarity: 68.06% Mismatches: 2
Query Match: 93.94% Indels: 495
DB: 6 Gaps: 3
US-09-721-114-2 (1-1057) x BD083700 (1-9455)
QY 1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
DB 4223 ATGGAGATTTGTCAGTAGATCAGAGGGAGCTCGTGTGTTGGGACGAACTGATGCTT 4282
QY 21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40

[illegible]

Qy	628	nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerIy	648
Db	7581	GCATGAGAGCGACGCTATTGACTCAGAGACTGATTGTTCTGCATCAACCGTATTCAATCCAA	7640
Qy	648	sThrThrAlaAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe	668
Db	7641	GACAACTGCTGATGATGATTGTATATAGTAGCTGCCAAGGATGGTTTCAGATTATGCATC	7700
Qy	668	rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl	688
Db	7701	AAAGTGTGTTTGACACATAATTCCCAACAGAAGTCCTTTGGCATCTCCAAAGTATACACAGAAGGA	7760
Qy	688	uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe	708
Db	7761	GTTTACAGGGTCATTTGGCATTGACCAACACAGAGTCTCCACATCTCCAGACCTTTTCAGTC	7820
Qy	708	rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe	728
Db	7821	TACTCAGGAACACGACACACATTTGGCATGGAAGAAATGGTCATATTGCTGCAAGCTC	7880
Qy	728	rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisThrGl	748
Db	7881	ACCACATATTTTCATCATCATGATCATCAGTATATTGCTGAAGCACCACCTGAACATTTGGGG	7940
Qy	748	YArgLysAspAlaLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr	768
Db	7941	CCGTAAAGGCGCAAGACGTAAACGTTGGGAGCAATTTAAGGCCACTTACAAGAAATTTCTCC	8000
Qy	768	oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnIlaValAspLeuThrSerTh	788
Db	8001	AGACAGCAACATGTTGGTGTCTCAATTTTAGACCTGGTATCTCCAAGCAGATTGACTTGACTTCTAC	8060
Qy	788	rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs	808
Db	8061	TCATGTCATGGGATCTTCAGCCATTTATGCATCTGCCAACCCAGTAATTTGSCCAGTGA	8120
Qy	808	pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl	828
Db	8121	CCGCTATGCTGAAAGAGCGGTTAACCGGCTCCATGCAAGAAATTTTCCAAAGCACATAAGC	8180
Qy	828	aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr	848
Db	8181	AACCATGGAAGCGAGTAAGTTATGTGATCGGAATAATCTGCACAAAGTAGTCTTGTAATCC	8240
Qy	848	oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe	868
Db	8241	TAAAGAATCCATGCTCTGCCAGCGCATCTCTGAGAAATGATGGATCCATCAACATTAGCAAG	8300
Qy	868	rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl	888
Db	8301	CTTTCCCAACTATGGAACCTCTAGCAGGAACAGATGGAGTCTCAACTTCATATTTCTTCA	8360
Qy	888	nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl	908
Db	8361	GTATGCACATATCATCTAACAAGGATCAACAGCACATCATATGCAAGTAACCTCAATGG	8420
Qy	908	YLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr	928
Db	8421	AAAGATTCCATTTGACATTTGGAAGACTTTATCAGCGCATCAGCTGCGATGATCTGCACAGACC	8480
Qy	928	oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs	948
Db	8481	TTTACGCCCCACATCTTAGATTGGTGTGTTGGCTTCCTTGCCTGCAGAGAGAAATTTGCAAA	8540
Qy	948	nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIlePh	968
Db	8541	CTGGCTCGGACAACTGTGGGCACAAATCTGGTTATTAAGTTAGAGATGCTCAACAGGAATAAC	8600
Qy	968	rSerHisGlnMetAsnArgGlyHisPheGluAlaLeuAsnSerGlyMetPheSerAl	988
Db	8601	ATCGCATCATGATGAAACAGAAAGGAACATTTGAAGCCCTGAATTTCTGGAAATTTTTCAGC	8660
Qy	988	aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr	1008

/db_xref="GI:7630238"
/translaton="MPSMVDRAATSDMLIGPWAKNNISICDINCRDPGQSKOVVKAAL
KRIGKPKVQVIALTLETAIKNCGDI FHMVAERDVLHEMVKI VKKSDQVKKVKY
LTDITWQEAAGFPAPYQYAAHDLVRAAGAFKPSRDPAPLFGQSPAGRNWRS
PDQDEASAGANDFPALSSSEIQNARGIMDLVAEMNALDPNGREGLREVETVLVD
QCTYKGRVVLNNATBELMSGGLANDLQRLAKHDIAIAGLAIVRVEKPKSLQ
ALVDTBMQDCKKQALVDIEDPTQETNKPSQASVQSPFEOALAVPEPPVNGS
ATAPKADGVLLSSDDNFTSSNLSALVPTDPVVDSTPSONALAIVDIPSQNNT
NNTAKPADPQGVNSSTLPQSGQTPNATQPILOAQOQOQVGLPWGGAWNPCTSYPT
SSHGNGQIANNAPPAQQTVNYGIAIYIYISQPSYFTIPIIGCNHACPLCSV"
30166..30675
/notes="unnamed protein product; ESTs
AU031545 (B61825), AU031546 (B61825) correspond to a region
of the predicted gene.
hypothetical protein"
/codon_start=1
/protein_id="BA094772.1"
/db_xref="GI:7630238"
/translaton="MYNQPGAMQPSQPAVTQSGFANQPGSMPPHSPMCPQPPGM
QPSMPGAQPMYMAFGMPGMAFAMPQPMYQPMQSYQLVQQQAAQYSSQGR
PTYTAGNDISQKMYGLSMQDSSYPTSTPSSSSMGQSPDKFLGDLSS
IAKTQKNA"
complement [join (34629..34795,36831..36867,38803..38895,
41780..41830)]
/notes="unnamed protein product; hypothetical protein"
/codon_start=1
/protein_id="BAA94773.1"
/db_xref="GI:7630240"
/translaton="MTVVRCTTQRGLLEGRICQDIRGKILSGSKPFDSPKRNKGTGI
EIEVGLDACYVITVTKPENERRIKKHGRIGLSRLCDRLGPKLVWDPSTKVVTPY
VQPRGHTLPHRG"
join (46367..46423,47537..47769,48031..48153,50082..50299,
51762..51864,52091..52204,52433..52532,52784..53004,
54084..54514,54601..55709,56020..57468)
/notes="unnamed protein product; hypothetical protein"
/codon_start=1
/protein_id="BAA94774.1"
/db_xref="GI:7630241"
/translaton="MPNGLIKRSEELDKLIAHATPHPPPPQIFVGAAAAATATA
WICAPVAAAMICAPVADQSPGSCLRVAEPNNRQTPRPSVVAESPDIHRHPILL
IGCWKDKILFTWGHGAVVADQSVRSRARTKIRGVGLMVDGYPPIHVGVVG
PMCTAVLHRMSGNGRLALITCSITKANCILAMKLELPEKQTFELGNIEBAASG
ERSTARRVDRAPLPRRRCRGRKPPSCAPLSGHELLASMGTKALI FVHAEKKIPF
AGILAAWMLFLAIILPARISGALIPHELIYDVIYACLVLCVLPVDMISVAV
DQGAUVGTGNMARGTGAVALIELATPDRQAAEAAGVDPAQHCHEFSIRGY
VALLQKDPKFCISLIRPHDQKQDEHKASSPFAKPRWDCSKCLDKLTSDNQT
APRTLPAQKQTSQCSITFVRSITFVPAVSQSVKSTQSGKNADSTLPKSVQE
GNDSKCAPSGKGAATDS PMKDLQGAQNDVAANTSDNITKDPNPMGSKERQD
TWHLVNCADQPPSPKLSVVLXNRDENGKRETLVABQCNITKDPNPMGSKERQD
VAEQNLTKDPKPSYQKQCEONEPBEVVLKPSKSKRKTDKLKKQKSKRTTA
QADVSADKLKPKKPKVLLSEIINAQVENSRSDEVHRENAADPCEDRSTIPVPMK
VSMIDIPENHTVTGSDGLSKSKNKTXYSDVDGSSLMNLINGKKRSTSVHHTVA
PAGMLNSKKTVPYTTAQHDDENGTLDTNMHTDQVHSEISTQRCSKSGTAGL
SKGTHSAASTKYGESSTQNTIHLVSAEQCMETNSVLHSKAVSAEHDQIMR
SLDKHQSLPKKKKQLEVTREKQITDDIIPMDIVELLAKNKHQELMTDQSDINR
IQSTKTDADDQYVVAADQSDVYASPLFSHDDQYIAEAPTEHGRKADKALTWEQF
PQNCSTQOQTHLRMEBMTVIAASSPLFSHDDQYIAEAPTEHGRKADKALTWEQF
KATTNSPAAQCGQPPRIQAVDLTSHWSSNSYASRQPVLAIDRAVAERANOV
HARNPSTIAWESKJCDRRNAQVVLIPKESNPETHLHMMDPSTILASFPFYGTSS
RNQMESQLHNSQYAHNQKGTSTSYGNLNGKIPLETEDLSRHLQHLDRPLRPHKR
VGLVLSLQKSIANWENSCQYKSGKGLGVSTGITSQHNKKEHFEELPFAWNSA
LQLGSVSSSADPLTSLARNSIAQSWTRGKGMVYHDFRVQRTITNKNPADFTTSSND
NEMVDR"
complement [join (58634..58667,58796..58902,59018..59083,
59368..59432,59585..59648,60333..60442,60620..60707,
60835..60945)]
/notes="unnamed protein product; ESTs
AU082698 (R0665), D32963 (R0665) correspond to a region of
the predicted gene.
Similar to Arabidopsis thaliana rac GTP binding protein
Arac7. (AF079484)"

Alignment Scores:

[illegible]

55938	Db	ACGTGCGTGGCGCAATATTCTTTTTTTAGACTCATATATATAGTGAATTGTAATAGCATGTGACA	55999
576	Qy	-----ValSerProAlaGluHisAspIleGlnIleMetSerAs	588
55998	Db	TTTTCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA	56057
588	Qy	pLeuHisGluGlnSerIeuProIlyssIlyssIlyssGlnIlyssLeuGluValThrArgG	608
56058	Db	CTTTCTATGACGAGAGCTTACCCCAAGAAGAAAGAGCAAAACCTTGAAGTGCATCTCGGA	56117
608	Qy	uIlyssGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnG	628
56118	Db	AAACACAGACCATGATAGATGACATCCCATGATATTGTTGAACCTGTAGCTAAAAACCA	56177
628	Qy	nHisGluArgGlnLeuMetGluThrAspCysSerAspIleAsnArgIleGlnSerIy	648
56178	Db	GCATGAGAGCAGCTATTATGACTGAGACTGATTTGTTCTGACATCAACCGTATTCAATCCAA	56237
648	Qy	sThrThrAlaAspAspAspCysValIleValalalalauysAspGlySerAspTyrAlaSe	668
56238	Db	GACAACTGCTGATGATGATTTGTTAATAGTAGTCTGCCAAGGATGGTTTCAATTATGCATC	56297
668	Qy	rSerValPheAspThrAsnSerGlnGlnIlyssSerIeuAlaSerGlnSerThrGlnIlysg	688
56298	Db	AAGTGTGTTTGCACATTAATTTCCCAACAGAAAGTCCTTGGCATCCCAAGATGACAGAAAGCA	56357
688	Qy	uIeuGlnGlyHisLeuAlaIeuThrThrGlnGlnSerProHisProGlnAsnPheGlnSe	708
56358	Db	GTTACAGGGTCAATTTGGCATTTGACACACAGAGTCTCCACATCTCCAGAACTTTTCAGTC	56417
708	Qy	rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe	728
56418	Db	TACTCAGGAACACAGACACACATTTGCGATGAGAAATGGTCACTATTGTGCAAGCTC	56477
728	Qy	rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTyrG	748
56478	Db	ACCACATATTTTTCATCATCATGATGATCAGTATATTCTGAAGCACCACACTGGAACATTTGGGG	56537
748	Qy	YArgIlyAspAlaLysIlyssLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr	768
56538	Db	CCGTAGGAGCGCAAGAAAGCTTAACGTGGAGCAATTTAAGGCCACTACAGAAATTTCTCC	56597
768	Qy	oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh	788
56598	Db	AGCAGAACATGTGGTGGTCAATTTAGACTCGGTATCCAAAGCAGTGTGACTGTACTCTAC	56657
788	Qy	rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs	808
56658	Db	TCATGTCTGGGATCTTCAGCAATATATGCATCTGCCAACCGTAAATTGGCCACTGGA	56717
808	Qy	pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl	828
56718	Db	CCGCTATGTCTGAAGAGCGGTTAAACAGGTCCATGCAAGAAATTTTCCAAACCAATAGC	56777
828	Qy	aThrMetGluAlaSerIlyssLeuCysAspArgAsnAlaGlyGlnValValIleuTyrPr	848
56778	Db	AAACATGGAAGCGAGTAAATGTTATGTATCGGAGAAATGCTGGACAAGTAGTCTTTGTATCC	56837
848	Qy	oIlyssGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe	868
56838	Db	TAAAGATTCATGCCCTGCGCGGCATCTCTGAGATGATGATTCATCAACATTAGCAAG	56897
868	Qy	rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG	888
56898	Db	CTTTCCCAACTATGGAACCTTCTAGCAGGAACAGATGGAGTCTCAACTTCTAATAATCTCA	56957
888	Qy	nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG	908
56958	Db	GTATGCACATTAATCAGTCAAAAGGATCAACCAAGCATCTCATATGCGCATGTAACCTGAAATGG	57017
908	Qy	yIlyssIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr	928
57018	Db	AAAGATTTCAATTCACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACACACC	57077

Qy	928	0leuAArgp-ohISProArgValgVlValleuGlySerLeuLeuGlnlySgllleAlaAs 948
Db	57078	TTTACGCCCACTCTCTAGAGTGGTGTGCTTGGCTCTCTGCGAAGAGGAAATGCGAAA 57133
Qy	948	nTPSPserGluAenCysGlyThrGlnSerGlyTrlyrlystLeuGlyValserThrGlyIleTh 968
Db	57138	CTGGTCGAGAACTGTGGACACAACTCTGTTTATTAAGTTAGAGTGTCAACAGGATTAAC 57197
Qy	968	rSerHisGlnMetAsnArgLySgLUHISpHeGluAlaLeuAenSerglyMetPheSerAl 988
Db	57198	ATCGCATCAGATGAAACAGAAAGCAACATTTTGAAGCCCTGAAATCTGGAATGTTTCAGC 57257
Qy	988	alysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
Db	57258	AAATGAATGATCATTCAGTCTGGTCTCTGTAGCTCCAGTCAGATTTTTTATCAGGAG 57317
Qy	1008	gAsnSerIleAlaGlnSerTrpThrArgGlyLySgLySgLyMetValHisProLeuAspAr 1028
Db	57318	GAACAGCATAGTCTCAATCTTGACACAGAGGCAAGGTAATAATGGTTTCATCCCTTGGATCG 57377
Qy	1028	gPheValArgGlnAspIleCyleThrAsnlyAsnProAlaAspPheThrThrIleSe 1048
Db	57378	GTTTGTGACAGCAGGATCTGTATTAACATAACAGAACCCAGCTGATTTTATCTACATCAG 57437
Qy	1048	rAsnAspAsnGluTyTrMetAspTyArg 1057
Db	57438	TAAACGATAACGAGTATATGATTACCGC 57465
RESULT 5		
LOCUS	AKL03364	
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J03316EN23, full insert sequence.	linear PLN 24-JUL-2003
ACCESSION	AKL03364.1	GI:32988573
VERSION	AKL03364.1	
KEYWORDS	PLN CDNA; CAP trapper.	
SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group)	
REFERENCE	1	Agrobiological Sciences Rice Full-Length cDNA Project Team, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,X., Kishimoto,N., Yasaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otsu,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariwaka,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Tsudomegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Otsu,N., Oka,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
AUTHORS	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice	
JOURNAL	Science 301 (5631), 376-379 (2003)	
MEDLINE	22752273	
PUBMED	12869764	
REFERENCE	2	(bases 1 to 3626)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,X., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kawai,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,	

Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kuroesaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, Y., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, Q., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-338-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodana, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, E., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirose, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

source
1. 3626
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033126N23"

ORIGIN
Alignment Scores:
Pred. No.: 1,696-302 Length: 3626
Score: 5179.00 Matches: 992
Percent Similarity: 99.90% Conservatives: 0
Best Local Similarity: 99.90% Mismatches: 1
Query Match: 93.72% Indels: 0
DB: 8 Gaps: 0

US-09-721-114-2 (1-1057) x AK103364 (1-3626)

Db	707	GCTCGTGGTGAACCTGGTGTCTAGCCCGCAGTGTGGAGCTGACAGCGACCCCTCGTCTGAG	766
Qy	41	AspAlaAlaAlaGluAlaGluValAspGluProAlaGlnHisGlnCysGluHisPheSer	60
Db	767	GATGAGCCGCTGAAGCTGGTGTAGACGACCGGCACACACACCATCGAGCAATTTCTCC	826
Qy	61	IleArgGlyTyrValAlaLeuLeuGlnLysIleAspProLysPheCysSerIleSerArg	80
Db	827	ATAAGAGGGTATGTGTCTCTCTTTCAGAGAAGATCCAAATTTCTGCTCTCTCTCTCGG	886
Qy	81	IlePheHisAspGlnLysLysCysAspGluHisLysLysLysSerSerSerProPheSerVal	100
Db	887	ATTTTCATGACGAGAAAATGATGTGACACAAAGCTAGTTCAAGCCCAATTTCTGTA	946
Qy	101	AlaLysPheArgArgTTPAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn	120
Db	947	GCAAGATTTCCAGCATGGATTTGCTCGAAGTCTGTGGATAAGTTGAAACTTCAGATAAT	1006
Qy	121	GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle	140
Db	1007	GGAACGACCAACAAAGACTCTTCCCGCAAGCAAGATGSCAACAGATGATGGTGTCTCCATC	1066
Qy	141	ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer	160
Db	1067	ACATTTGTTTCGAGCACATTTTGTGCTGCTAGTGTGGTTCCTCAAAAAGTGTCTCTAGC	1126
Qy	161	ThrGlnSerSerGlnGlyLysAsnAlaLysArgSerThrLeuProLysSerValGlnGlu	180
Db	1127	ACACAATCATCTCAAGGAGAAATGCTAGATCACTCTTCCAAAGAGTGTGCAGAA	1186
Qy	181	GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr	200
Db	1187	GCGAATGACTCCAATGCAATGGGCTTCTGGCAAGATGGAGCTGCTGAGGCCAATACT	1246
Qy	201	AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal	220
Db	1247	GATTCACCAATGAAAGATTTGCAAGGCGCAGCCCAAAATTTATGATGTGGCAGCAATGTC	1306
Qy	221	SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp	240
Db	1307	TCTGAGGACACACATCTCTGTGATGTTGGGCTTTACTGAGTTCCTCCAGATTCATCG	1366
Qy	241	HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal	260
Db	1367	CACATAGAAGTAATGTTGTCGATCAACCTCCATCCACTCCCAAAATTTCTGNAAGTGTTC	1426
Qy	261	LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys	280
Db	1427	CTCAAAAGAAATGAAAGATGAAATGAAATGAAATGAAAGAGACTCTTGTGCTGAGCAGTGC	1486
Qy	281	AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln	300
Db	1487	AATTTGACCAAGATCTTAACCCCAATGCTCGAAAGGAACGTGATCAGGTTGCTGAGCAG	1546
Qy	301	CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn	320
Db	1547	TGCAATTTGACCAAGATCCGAAACCAAGTGTCTGAGGAGAAATGAGCAGATCTGCAT	1606
Qy	321	GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys	340
Db	1607	GAGCCATGTGAAGAGTTGTTCTCAAAAGAGCTCCAAATCTAAGAGGAGACGAGATAAG	1666
Qy	341	LysLeuMetLysLysGlnGlnHisSerLysArgThrAlaGlnAlaAspValSerAsp	360
Db	1667	AAGTTGATGAAGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAT	1726
Qy	361	AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla	380
Db	1727	GCAAGACTTTTTCGAGGAGAAAGCCAAAGAGTGTGGCTTCTATCAGAAATTAATAATGCT	1786
Qy	381	AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys	400
Db	1787	AACCAGGTTGAGGATTTCTAGAAGTGACGAGTTTCATCTGCGTGAATGCCGCTGATCCCTGT	1846

QY 401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
 DB 1847 GAGGATGATGAAGTACCATCCCGTCCCGATGGAGTAGCATGATATCTCTGTAGC 1906
 QY 421 AsnHisThrValGlyGluAspGlyLeuLysSerSerIysAsnLysThrIysArgLysPyr 440
 DB 1907 AACCATACAGTGGGAGAGATGGTTAAATCAAGTAAGAACCAAGACAAACGCAATATC 1966
 QY 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysVal 460
 DB 1967 TCGATGTTGTAGATGATGATCATCTTATGACTGGCTGATGGAATGGAAAGAAAGAGA 2026
 QY 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
 DB 2027 ACTGGAAGTGTGCATCACACAGTGTCTCTCCAGTGGGAATTTGAGCAACAAAAAGTG 2086
 QY 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
 DB 2087 ACACCCACTCGGAGTACTCGAGCATGATGATGAGAAATGATGATAAATGGCTTGTGACACA 2146
 QY 501 AsnMetHisLysThrAspValCysGlnHisValSerGluLysSerThrGlnArgCysSer 520
 DB 2147 AATATGATGAACAGATGTCTGTCCAGCATGTATCAGAAATCTCCACAGAGGTGTCTCA 2206
 QY 521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
 DB 2207 TCAAGGGGAAAAACAGCGGTTTGGTAAAGGGGAAAAACACATTCAGCTGTGTAGTACCAAA 2266
 QY 541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
 DB 2267 TATGGTGGTGAAGCACAGAAATGTGTACAGATATGATGATGATGATGATGATGATGATGAT 2326
 QY 561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
 DB 2327 TGCAGATGAAACCGAAAACTGTCTCAGTCACTCGGCAAAAGTGTTCCTCCAGCTGAG 2386
 QY 581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
 DB 2387 CATGATATCCAAATATGTCTGACCTTCATGAGCGAGGTCTACCCAGAAAGAAAGAGAG 2446
 QY 601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
 DB 2447 CAAAACTTGAAGTGACTCGTGAAGAAACAGACCATGATGATGATGATGATGATGATGATGAT 2506
 QY 621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
 DB 2507 GTTGAACCTGTAGCTAAAGAACCCAGCATGAGAGGCGAGCTTATGACTGAGACTGATGTCT 2566
 QY 641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
 DB 2567 GACATCAACCGTATTCAATCCAGACCACTGCTGATGATGATGATGATGATGATGATGATGAT 2626
 QY 661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
 DB 2627 AAGGATGGTTTCAGATTATGATATCAAGTGTGTGTGACACATTAATCCCAACGAGATCTCT 2686
 QY 681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
 DB 2687 GCATCCCAAGTACAGAGAGAGTATACGGGTCTATTGGCATTTGACCCACACAGAGTCT 2746
 QY 701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuArgMetGluGlu 720
 DB 2747 CCACATCTCAGAACTTTTCAGTCTACTCAGAACAGCAGACACATTTGCGGATGGAAGAA 2806
 QY 721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
 DB 2807 ATGGTCACTATTGTCTGCAAGCTCCACTATTTCATCATCATGATGATGATGATGATGATGAT 2866
 QY 741 GluAlaProThrGluHisThrGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
 DB 2867 GAAGCACCACTGACATTTGGGCGCTAAGGCGCAAGAGAGTAACTGTGGGACCAATTT 2926

QY 761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
 DB 2927 AAGGCCACTTACAAAGAAATTTCTCCAGCAGCAACATATGGTGTCTCAATTTAGACCTGTATC 2986
 QY 781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTrpAlaSerArg 800
 DB 2987 CAAGCAGTTGACTTGCATCTTCTACTCATGTCTGGGATCTTCAGCAATATTATGCACTCTCC 3046
 QY 801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
 DB 3047 CAACCAAGTAATTTGCCCACTGCGCCATGCTGTAAGAGAGCGTTTAAACCAAGTCTCCATGCA 3106
 QY 821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLysCysAspAspArgAsn 840
 DB 3107 AGAAAAATTTCCCAAGCACATAGCAACCACTGGGAAGCGAGTAAGTTATGTATGATCGGAAAT 3166
 QY 841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
 DB 3167 GCTGGACAGTAGTCTTGTATCTTAAAGATCTCATGCTGCGACGCACTCTCTCGAATG 3226
 QY 861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
 DB 3227 ATGGATCCATCAACATTAGCAAGCTTCCCAACTATGGAACTTCTAGCAGGAACCAAGATG 3286
 QY 881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
 DB 3287 GAGTCTCAACTCATTAATTTCTCAGTATGCAATATATCATGATCAAGGATCAACAGCACA 3346
 QY 901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
 DB 3347 TCATATGGCAGTAACCTGAATGGAAGATTTCCATTGCATTCGAGACTTATCACGCGCAT 3406
 QY 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
 DB 3407 CAGCTGCATGATCTGCACAGACCTTTTACGCCACATCTCTAGAGTTGGTGTGTGTGCTCC 3466
 QY 941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
 DB 3467 TTGCTGCAGAGAGAAATTCGAACTGGTGGGAGAACTGTGGCACAACAATCTGCTTATAAG 3526
 QY 961 LeuGlyValSerThrGlyLysThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
 DB 3527 TTAGGAGTGTCAACAGGAATAACATCGCATCATGATGAAACAGAAAGGAACATTTTGAGCC 3586
 QY 981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeu 993
 DB 3587 CTGAATCTTGGATGTTTTCAGCAAAATGGAATGCATTG 3625

RESULT 6
 AX590041
 LOCUS 3291 bp DNA linear PAT 24-JAN-2003
 DEFINITION Sequence 223 from Patent WO02081695.
 ACCESSION AX590041
 VERSION AX590041.1 GI:27901181
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1.
 AUTHORS Zhu, T., Glazov, E.A., Meins, F., Wang, X. and Chang, H.S.
 TITLE Genes that are modulated by posttranscriptional gene silencing
 JOURNAL Patent: WO 02081695-A 223 17-OCT-2002;
 Syngenta Participations AG (CH)
 FEATURES
 Location/Qualifiers
 1..3291
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Alignment Scores:

Pred. No.: 2,32e-14 Length: 3291
Score: 393.00 Matches: 248
Percent Similarity: 36.83% Conservative: 191
Best Local Similarity: 20.81% Mismatches: 401
Query Match: 7.11% Indels: 352
DB: 6 Gaps: 58

US-0721-114-2 (1-1057) x AX590041 (1-3291)

QY 55 GlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysAspProLys 74
DB 73 AAATGTCATCTTTCCATGCGTGGATTCGTAGCTGAACTCGTGAGAGAGACCTTAGA 132
QY 75 PheCys-----SerLeuSerArgIlePheHisaspGlnLysCys 88
DB 133 AAATGTCGCGGTTTTCAGAGAGAGTGTAGTTAGTA----- 171
QY 89 AspGluHisLysAlaSerProPheSerValAlaLysPheArgIlePheArgPheCys 108
DB 172 GACCAACAAGCTATCTCTCTACTTATCTGTTCCAAAGTTAGATGGTGGCATGT 231
QY 109 SerLysCysLeuAspLysLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
DB 232 ATGAGCTGCATCAAGATATA-----GATGCTCATGGGCCAAAGATTTGTGGA 279
QY 129 AlaLysGluAsnGlyThrSerAspGly---CysSerIleThrPheValArgSerThrPhe 147
DB 280 CTGCATTCAAACTCAAAGCTATGGAAGTCTCTCTGTTATCGAAGTAAAGCAAGTTC 339
QY 148 ValProAlaSerValGlySerGlnLysValSerProSerThr----- 161
DB 340 AATTCGCTAACTATCATTCATGACGAGAGAAAGAAAGAACTGATATTGCAGATAATGCT 399
QY 162 ---GlnSerSerGlnGlyLysAsnAlaAspArg-----SerThrLeuPro 175
DB 400 ATTGAGGAGAAAGTGGGTGAAACTGTGAGATGATGATGATGATGATGATGATGATGAT 459
QY 176 LysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAla 195
DB 460 AAGAAGCAGCT-----GGTCGACCTATGGT----- 486
QY 196 AlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAsp 215
DB 487 ---GCTTCTAATGTAGGCAAGACAGACAGAAAGCTGTGAGTCCGAGCAG----- 534
QY 216 ValAlaAlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeu 232
DB 535 GTAGGAACAACAGATCTAAAGAAACAAACAAACCATCGATGATGATGATGATGATGATG 594
QY 233 ProGluValProGlnIleThrPheIleGluValAsnGlyAla----- 247
DB 595 AAAGAGAAACAAATGTGGATCAGGCTGTGACAAACGTTCCGCTCATCTGAAATTCGTGT 654
QY 248 -----AspGlnProSerThrProLysLeuSerGluValValLeuLysArgAsn 264
DB 655 GTGTTGAGGATACACCCCTTAGGCAACCAAGATCATAAAGGCATTCGCGTCTGATG 714
QY 265 GluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLys 284
DB 715 GAATGCGTAACCGGTCATCAGAAAGTAAATCTGCT----- 753
QY 285 AspProAsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCys 301
DB 754 -----ATGATGGGTTCAGCGTAGGAATCTCGCAGAGTTCTCTACTCAGT 801
QY 302 AsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGlu 321
DB 802 GAGTTGCTGGTAATACAAACCAACGAGTGGTAGTAACATCAGAAAA----- 849
QY 322 ProCysGluGluValValLysArgSerSerLysSerLysArgLysThrAspLysLys 341
DB 850 -----GAAGAGTCTGCTTTGAAGAGGAATCAGTTAGAGGTGAAAA-----AGAAAG 897

QY 342 LeuMetLysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp 360
DB 898 TTGTTACTGAAACAAATATGTCAGCGCGGATATTAGTATCAATGGGTGCAACCTCTGAA 957
QY 361 -----AlaLysLeuCys----- 364
DB 958 AATGCTTCCAAAAGTTGTGACTGTGATCAAGGTAAATAGTCAATCACTGATAGTGGGTT 1017
QY 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluLeu 377
DB 1018 GACAGAACTCCATTTAAAGGTAAGCAGAGAAACAGAGATTTTCAAGTTGTTGACGAGTT 1077
QY 378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
DB 1078 GTACCATCAGTCTCTCTGTGAAATCTCAA-----GAAGGTATCATCAGGAGCATGATCA 1131
QY 398 AspProCysGluAspAspArgSerThr----- 406
DB 1132 GATCTCTAGTAAG-----AGATCAACTCTCTCGGCACTTTTATTCCTGGAACCACTTCT 1185
QY 407 IleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGlu 426
DB 1186 GTTCTCTGTCCT-----COGGGTACTCAGAGAACA-----CAG 1218
QY 427 AspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp 446
DB 1219 AGGAAGCTCAGTTTACCACAGAGAGACAAAGAG-----CCTGTATCGATAAT 1269
QY 447 GlySerSerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArg 460
DB 1270 GGAAGAGCACTGTGATCAGTTTAGTAACCGCATTTGATGAAGTCAAGTTAACTCGCAT 1329
QY 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysVal 480
DB 1330 ACTGTCCTCCATCGAACACACAGATATCCAAACTCGAGACTTATTTGAATGGGAAAGG 1389
QY 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
DB 1390 -----GGCGGTTTATTTCAC 1404
QY 501 AsnMetHisLysThrAsp-----ValCysGlnHisValSerGluIleSerThrGlnArg 518
DB 1405 AACCGTTTGGCTTCAGATGGATATTTTCAGAAATATCTCTCTCAGGTTAATGATAAGCG 1464
QY 519 CysSerSer-----LysGlyLysThrAlaGlyLeuSer 529
DB 1465 ATAACATCTTGGCATTTGCAAGACAAATGATATGAGGTCAAGAGACGCGGAACCAAC 1524
QY 530 LysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGly 549
DB 1525 TGTCTTCGAGATTTTAGTCTCTCTTAAATCCAGCTCAGGTGGATGTTGAGAACTGGA 1584
QY 550 GlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerVal 569
DB 1585 -----GTAGATATTGTTGACTTCAGAAACCAACCAACCAATACAAACAGATCGTCTTC 1638
QY 570 LeuSerHisSerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeu 589
DB 1639 TCGAACTTGAAGCTAAGATACCCCTTCTCTACTGAGTTGCGGATTTATCTCGGTG 1698
QY 590 -----HisGlu 591
DB 1699 CTGCAAAAAGAGTCTTCTGTCAGATAGAAAGGGAAGACTGTTTATGTCACAAACAT 1758
QY 592 GlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThr 611
DB 1759 CATGAGACCAAGAGAGCCAAAGTCCAGATAGAAAGGAGACTACCACTGAAGAGCAAAAC 1818
QY 612 MetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArg 631
DB 1819 AAC---GATGATATTCCAAATGAGATAGTGGAGCTCATGTGCCCAAAACCAAGTACGAGAG 1875
QY 632 GlnLeuMet---ThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysThrThr 650

1876	DB	1876	FTCTTTCCCGCAAAAGAGAGATGTTAGCAACAAACAGCCATCACAAAGAAACAGACAC	1935
651	QY	651	AlaAspAspCysValIleValAlaAlaIleAspGlySerAspTyrAlaSerSerVal	670
1936	DB	1936	AAATCCAGAAATGCTCTACTGATGTATCTCAATGAACCTACGATACCGGATCTCACTT	1995
671	QY	671	PheAspThrAsnSerGlnGln-----LysSerLeuAlaSerGlnSerThrGlnLysGlu	688
1996	DB	1996	GAGGACACACACATCAAGACCACCAAAACCGGTAGTAGCAACCAAGAGGAGGAGAA	2055
689	QY	689	LeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHis-----	702
2056	DB	2056	-----CAITTTCTATGGGAGACAGACAGACTCTCATGACTTCTTCCCAATAAGT	2106
703	QY	703	-----ProGlnAsn-----PheGlnSerThrGlnGlnGlnThrHis	715
2107	DB	2107	CAGCCTTATGTCCTCTCGTTTGGGATCTTCTCTCTACCCCAAGAAAACCGA-----	2160
716	QY	716	LeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAsp	735
2161	DB	2161	-----GCGAGCTCCATCCGGTGTCTTCTGTGTCAACA	2190
736	QY	736	AspGlnTyrIleAlaGluAlaProThrGluHisThrGlyArgLysAspAlaLysLeu	755
2191	DB	2191	TGTCAGTGGCTTGGGAATTTGCCAATGTG-----GGTAATCAGAACCCTTCCCATCC	2244
756	QY	756	ThrTrrGluGlnPheIysAlaThrThrArgAsnSerProAlaIaIaThrCysGlyAla---	774
2245	DB	2245	TCATTTTGGGTAATTACGTGCT-----TGTCATACTTCCCAAGTGT	2286
775	QY	775	-----GlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGly	792
2287	DB	2287	CCTAATCAATACAGA-----GAAGCTTCTATCCAAATTGG	2322
793	QY	793	SerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaGlu	812
2323	DB	2323	CCATCTTCC-----ATGATACCACACAGAGTCAATACAAGCCA	2361
813	QY	813	ArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAla	832
2362	DB	2362	GTTCCTTTAAATTAATTAATAGTCAACAAAT-----CCGGTAGCTTTCA-----	2412
833	QY	833	SerLysLeuCysArgArgAsnAlaGlyGlnValValLeuTyrProLysGlu-----	850
2413	DB	2413	TCTAACAAATGAAATACATGAACCTCAACTTTGTTGCTGCCAACGGAAGCAAAATGT	2472
850	QY	850	-----	850
2473	DB	2473	GGCCCTAATCCAGAAATTTCAATTTGGCTGCAACATCTGCTGGGTTAGTAGTAGT	2532
851	QY	851	-----SerMetProAlaThrHisLeuLeuArg	859
2533	DB	2533	AGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTATACCGCATTTGCATCTACTCAGC	2592
860	QY	860	MetMetAspProSerThrLeuAlaSerPhePro-----AsnTyrGlyThrSer-----	875
2593	DB	2593	CTTCTGGATCCTCGCTGAGGTCAACGACTCCCGCTGACCAACACGGAACACATAAATTT	2652
876	QY	876	SerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLys	895
2653	DB	2653	ACTAAAGACATTTTCGCCAGCCAAACCTGTCACCAAGAGTTTATAGAGCTTCAACAGGG	2712
896	QY	896	GlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGlu	915
2713	DB	2713	GACTCTAGTAAGTCAGCCTACTCAACTAAG-----CAGATACCT-----TTTGAT	2757
916	QY	916	AspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeuArgPro---HisProArg	934
2758	DB	2758	CTATACAGCAAGAGATTCACACAGAGCCTTCCCGNAGAGTTTCCCATCTCACTCCACT	2817
935	QY	935	ValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrrpSerGluAsnCysGly	954

Db	2818	ATTGGAGC-----TCCTC	ATTTTCATTTTCATAATGCTTCATGGAGTCTCTCAT-----	2866
Qy	955	ThrGlnSerGlyTyrIysLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArg	974	
Db	2866		-----CATCAAGAGAGAGAAA	2880
Qy	975	LysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaIysTrpAsnAlaLeuGln	994	
Db	2881	ACCAAGAGAAAAGACACC-----TTTGCTCCTCTTTACATATCTCATGAA	2925	
Qy	995	LeuGlySerValSerSerSerAlaAsp-----PheLeuSerAlaAArgAsn	1009	
Db	2926	ARGSCGGTGTTCAGCAGCAGCAACCAAGCGAAGTTC	2985	
Qy	1010	SerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeu	1026	
Db	2986	TCC-----ATGATGCTTCTTGAAATTTTCAC	3012	
Qy	1027	-----AspArgPheValArgGlnAsp-----	1033	
Db	3013	ATGACGGATAAAGAAAAGAAACAAAGAGAAAAGACAGAGCTGCAATAACAAATGCTCT	3072	
Qy	1034		-----IleCysIleThrAsnLysAsnProAla	1042
Db	3073	GGCGGACCTGTGAAGAACAGTTCGACCCATTGTGTGCGAGCGTCAATGAACCCCTGCT	3132	
Qy	1043	AspPheThrThrIleSerAsnAspAsnGluTyrMet	1054	
Db	3133	GAITTCACCATCTCGAACCTGGGAATGTTTACATG	3168	
RESULT 7				
LOCUS	AR368885	8648 bp	DNA	linear
DEFINITION	Sequence 2 from patent US 6376751.			PAT 12-SEP-2003
ACCESSION	AR368885			
VERSION	AR368885.1	GI:34603269		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 8648)			
AUTHORS	Sung,Z.R., Aubert,D. and Chen,L.			
TITLE	Nucleic acids encoding EMF1 that control reproductive development in plants			
JOURNAL	Patent: US 6376751-A 23-APR-2002,			
FEATURES	Location/Qualifiers			
source	1. .8648			
ORIGIN	/organism="unknown"			
	/mol_type="genomic DNA"			
Alignment Scores:	4.07e-08	Length: 8648		
Pred. No.:	298.00	Matches: 225		
Score:	Percent Similarity: 35.53%	Conservative: 152		
Best Local Similarity:	21.21%	Mismatches: 360		
Query Match:	5.39%	Indels: 324		
DB:	6	Gaps: 50		
US-09-721-114-2 (1-1057) x AR368885 (1-8648)				
Qy	198	AlaAsnThrAspSerProMetIysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla	217	
Db	5090	TCTAATGTTAGGAGCAGACGAGAAAGCTTGTGAGTCCGAGCAG-----GTAGGA	5140	
Qy	218	AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu	234	
Db	5141	AACAAACAGTCTTAAGAAAACCTAAACAAACCATCGATGTGATATTAGCAGCTCGAAGAG	5200	
Qy	235	ValProGlnIleThrTrpHisIleGluValAsnGlyAla-----	247	
Db	5201	AAACAAATGTGATCAGCGTGTGACAAAGTTCGGCTCATCTCTGAAATTCGTGTGGTGT	5260	

QY 248 ---AspGlnProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Dd 5261 GAGGATACCACTAGGCAACCAAGCAATCATAAAGGCAATCGCGGTCTGATGATGC 5320
QY 267 GluAsnGlyLysThrGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Dd 5321 GATACCGGTATCAGAAAGTATATCTTGCT----- 5353
QY 287 AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu 303
Dd 5354 -----ATGAGTGGGTTCAGCGTACGAAATCTCGAAGGTCGTCTACTAGTGGTTG 5407
QY 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Dd 5408 CTGGTAATACAAAACCACTGGTGTAGTACATCAGAAA----- 5449
QY 324 GluGluValValLeuLysArgSerLysSerLysArgLysThrAspLysLysLeuMet 343
Dd 5450 GAAGACTCTGCTTTGAAGAAGGAATCAGTTAGAGTTCGAAAA-----AGAAGTTGTTA 5503
QY 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp----- 360
Dd 5504 CCTGAAACAATATGTCAGCGGATATTGAGTACAAATCGGTGCACACTCTGAAAATGCT 5563
QY 361 AlaLysLeuCys----- 364
Dd 5564 TCCAAAAGTTGTGACTCTGATCAAGGTAAATAGTGAATCAACTGATAGTGGTTTGACAGA 5623
QY 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluLeuLeuLeu 379
Dd 5624 ACTCCATTAAAGGTAAGCAGAGAAACGAAAGATTTTCAGGTTGTGACGAGTTGTACCA 5683
QY 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Dd 5684 TCACCTCCTGTGTAACCTTCACAA-----GAAGGTATCAAGGAGCATGATGCATCTCT 5737
QY 400 CysGluAspAspArgSerThr-----IlePro 408
Dd 5738 AGTAAG-----AGATCAACTCTCGCGCACTCTTTATTCACTGGAAACGATTTCTGTTCT 5791
QY 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Dd 5792 TGTCTT-----CCGGTACTACAGAAACA-----GAGAGGAAG 5824
QY 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Dd 5825 CTCAGTTTACCAAGAAAGAGACAAAGAA-----CCTGTAATCGATAATGGGAAG 5875
QY 449 SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly 462
Dd 5876 AGCAGTGTATCAGTTTGTAGTACCGCATTCGATGGAAGTCAAGTTAACTCCGCATCTGT 5935
QY 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Dd 5936 CTTTCCATGAACACAGATATCCCAACTCGAGACTTATTGAATGGGAAAGGTTGGCGGT 5995
QY 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp----- 499
Dd 5996 TTATTTGCAACACCGTTTGGCTTCAGATGGATATTTCCAGAAATATCTCTCAGGTTAAT 6055
QY 500 -----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluLysSer 515
Dd 6056 GATAAGCCGATACATCTTGTGATTCGACAGCAATGATTATGTCGAGGTCAAGAGACGG 6115
QY 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu 528
Dd 6116 GAACCAAACTGTCTTCGAGATTTTGTCTCTCTTCTTAAATCCAGCTCAGGTGGTGGTTG 6175
QY 529 SerLysGly-----LysThr 533
Dd 6176 AGAAGTGGAGTAGATATTGTTGATCTTCAGAAACAAACCAACATACAAACAGATGCTCT 6235
QY 534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553

Dd 6236 TTCGCAACTTGAAGCTAAGATACCCCTCTTCTTCTACTGAAGTTGCGATTTATCTCGG 6295
QY 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisLe 573
Dd 6296 GTGCTGCAAAAGGTACAGCTTTGTCATCTATTTTCTTCAATCGCCATGCTTTCTCATTC 6355
QY 573 ----- 573
Dd 6356 TGTTTTCAACTGACCTATCTTGATATATCCAAACTGGTGTCTTCACTGATGATGAT 6415
QY 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
Dd 6416 CAGGATGATAACGGGATGATTACTTGGCAGGATGCTTCTGTGTCAGATAGAAAGGGA 6475
QY 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Dd 6476 AGACTGTTATGCTCCAAAGACATCATGGAGCACCAGAAAGCAAGTCCAGATGAAGG 6535
QY 604 LuValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Dd 6536 AGACTACGACTGAGAGCAAAACAAAC---GATGATATTCCATGAGATAGTGGAGCTCA 6592
QY 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Dd 6593 TGGCCAAAACCACTACGAGAGGTGCTTCCCGCAAAAGAAAGATGTTAGCAACAAC 6652
QY 643 snArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspG 663
Dd 6653 AGCATCAACAGAAACAGCACACAAATCCAGATGCTCTACTGATGATCAATGAA 6712
QY 663 LysSerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA 681
Dd 6713 CCTACGATAACGGGATCTCACTTGAGGACAAACACATCAAGACACCACCAAAACCGTGA 6772
QY 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Dd 6773 GTAGCAACGCAAGGAGGAAGAA-----CATTTCTTATGGGAGACAGCAGACT 6823
QY 701 roHis-----ProGlnAsn-----PheGln 708
Dd 6824 CTCATGACTTCTCCATAAAGTCAGCCTTATGTGCTTCTCCGTTTGGGATCTTCTC 6883
QY 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaLysers 728
Dd 6884 CTACCAAGAAACCCGA-----GCCAGCT 6907
QY 728 erProLeuPheSerHisHisAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Dd 6908 CCATCCGGTTTCTGGTCACAACTGTCACTGCTGGTGGGAATTTGCCAATCTGT-----G 6961
QY 748 LysArgLysAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Dd 6962 GTAAATCAGAACCTTCTCCATCTCTATTCGGGTATTACGTGCT----- 7005
QY 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Dd 7006 --TGTGATCTTTCAGAGTGTCTCTAATCATACAGA----- 7041
QY 785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Dd 7042 --GAAGCTTCTCATCCCAATTTGGCCATCTCTCC-----ATGATAC 7078
QY 805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPhePro 825
Dd 7079 CACACAGAGTCAATACAGCCAGTCTTTTAAATATTAATCATCATCAAAAT---CCGG 7135
QY 825 erThrIleAlaThrMetGluAlaSerLysLeuCysArgAspArgAsnAlaGlyGlnVal 845
Dd 7136 GTACGCTTTCA-----CAGGCACTAACATGAATAATACATGGAACCTCACTTTGTG 7189
QY 845 alLeuTyrProLysGlu----- 850

Db 7190 CTGCCAACGGGAGCAAAATGTGGGCTTAATCCAGAAATTTTCATTTGGTGTGCAACATG 7249
Qy 851 -----SerM 852
Db 7250 CTGCTGGGGTTAGTAGTAGTAGTAGGCAATAGATAACTTTCTAGTGAGAGCTCTA 7309
Qy 852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro 870
Db 7310 TACGGGATTCATCTACTAGCTCTCTGATCTCTCGCTGAGTCAACAGCTCCCGCTG 7369
Qy 971 --AsnTyrGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 7370 ACCAACACGGGAACACTAAATTTACTAAAGACATTTTCCGCCAGCCACACGATCCAAAG 7429
Qy 888 lntYrAlaHisAsnGlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSerThrSer 908
Db 7430 AGTTATAGAGCTTCAACACAGGGGACTCTAGTAGTCACTGAGCTACTCACTAAG----- 7482
Qy 908 lylsileProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgp 928
Db 7483 --CAGATACCT-----TTTGATCTATACAGCAGAGATTCACACAGAGAGCTCCCGGA 7534
Qy 928 roLeuArgPro--HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluLeuA 947
Db 7535 AGATTTCCCATCATCCTCCACCTATGAGAG-----TCTTCACTTCAATTTCAAAATG 7588
Qy 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db 7589 CTTTCATGAGTCTCTCAT-----7605
Qy 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 7606 -----CATCAGAGAAGAAACCAAGAGAGAAAGACACC-----TTTG 7642
Qy 987 exAlaLysTyrAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 7643 CTCCTGTTTACATACTCATGAAAGCGGTGTTTGCACGAGCAACGACCAAGCGAAGT 7702
Qy 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db 7703 TCCAGCTGCTGGGAGCATCGAAATTC-----A 7729
Qy 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 7730 TGATGCTTCCTTTGAAATTCACATGACGGATGAAGAAGAAAGAAAGAAAGAAAGAACG 7789
Qy 1034 -----IleC 1035
Db 7790 AGAGCTGCAATPACATGCTCTCGGGACCTGTGAAGAACAGTTCTGGACCCATTTGTT 7849
Qy 1035 ysileThrAsnLysAsnProHisAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db 7850 GCAGCGTCAATAGAAACCTGCTGATTCACCATTCCTGAACTGGAACCTGGGAATGTTTACATG 7908
RESULT 8
AF319968 8648 bp DNA linear PLN 05-SEP-2001
LOCUS Arabidopsis thaliana embryonic flower 1 (EMF1) gene, complete cds.
DEFINITION AF319968
ACCESSION AF319968.1 GI:15430696
VERSION
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
1 (bases 1 to 8648)
Aubert,D., Chen,L., Moon,Y.H., Martin,D., Castle,L.A., Yang,C.H.
and Sung,Z.R.
REFERENCE EMF1, a novel protein involved in the control of shoot architecture
AUTHORS and flowering in Arabidopsis
JOURNAL Plant Cell 13 (8), 1865-1875 (2001)
MEDLINE 21380417

11487698
2 (bases 1 to 8648)
REFERENCE Rubert D., Chen, L. and Sung, Z.R.
AUTHORS Direct Submission
TITLE Submitted (07-NOV-2000) Plant and Biochemistry, Riscoe,
JOURNAL Frederiksbergvej, Roskilde 4000, Denmark
FEATURES
Location/Qualifiers
1..8648
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
1..8648
/gene="EMF1"
1..3201
/gene="EMF1"
3202..3265,4160..4335,4448..4623,4704..4823,
4903..4956,5046..6307,6448..8648
/gene="EMF1"
/product="embryonic flower 1"
join(4241..4335,4448..4623,4704..4823,4903..4956,
5046..6307,6448..8031)
/gene="EMF1"
/codon_start=1
/product="embryonic flower 1"
/protein_id="AAK98528.1"
/db_xref="GI:15430697"
/translation="MGSSIKINSISIDLAGAANEIDMVKDHFSMRGFAETREDLR
KCPSEESVSVLQDQSYTLPTLSVFKRWCHMSICIDIDAHGPPDCGLHNSKAIG
NSVIVSKSFLNLTIDHEKKTIDADNAIEKVGVCNENDDQATTFKCARGP
MGASVRSKSRXIVSPQVGNRSKELANPSMDISSWKEKONVQAVDTTFGSSBIAG
VVEDTPPKATKHGIRGLMECDNSESINLMSGLORRKRKRVLLSELGLTKTS
GGSINRSALKESVRGKRLILPENNYVSRILSTMGATSENASKSCDSQGNSES
TSGSDRTDFGKQRNRQVDFVPSLFCISQEGIKHDADPKRSTPAHSLFTG
NDSVPPQPTQTERKLSPKKTKPVINDKSTVSPNSGIDGQVNSHTGSPWNT
VQTRDLNGKRVGLFDNRLASDGYRXYLSQVNDKPTLSLHQNQDYSRSDAPRN
CLURDSKSSSSGWLRTGVIVDPFRNNHNTNRSSPSNKLRLVPPSSTEVADLRV
LOKDSAGDKGTVQSHQAPRSQHDRETTTEBONDDIPWEIVELAKNOYE
RGLPKEEVSNKQSPQETAKSKNALIDINETYDNGISLENNITSRPKPCSSNAR
RHEHFMPSQKSHDFPPIQPVSPSPFGLIPEPTQENRASSIEFGSCNOMLNLPTV
GNQSPSPFRVLACDTQCVQPNQREASHFIMESSMIPPOQYKPVSLINQSTNP
GTLQASNNNTNLFVAAKGKQGNPFESFGCHAAQVSSSSSRSDIDNPFSSSS
ITPALHLLSLDPLRSTTPADQKNTKTRHPFPAQSKETIELQDSSKSAYSTK
OIPFDLYSKRFTQEPKSPPTTPICTSSLSQNASWSPHFKETKTKRQDPAPYN
THEKPVFASNDQAKFLQAGSNMPLKFTMTDKKQKQKABSCNNNAGPVKN
SSGPVCSVNRNPADFTIPEPGNVYMLTGEHLKVRKRTTKKKPAVKQDAMKQTKPK
VCPPTQNA"

Alignment Scores:
Pred. No.: 4,07e-08 Length: 8648
Score: 298.00 Matches: 225
Percent Similarity: 35.53% Conservative: 152
Best Local Similarity: 21.21% Mismatches: 360
Query Match: 5.39% Indels: 324
DB: 8 Gaps: 50

US-09-721-114-2 (1-1057) x AP319968 (1-8648)

Qy 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db 5090 TCTAATGTTAGGAGCAAGACAGCAAGATGTTGAGTCCGAGCAG-----GTAGGA 5140
Qy 218 AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu 234
Db 5141 AACACAGATCTAAAGAAAACCTAAACAAACCATCATGATGATATTAGCAGCTGGAAGAG 5200
Qy 235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla----- 247
Db 5201 AAACAAAATGTCAGCTGAGCAGACACTTCGGCTCATCTCGGCTCATCTGAATTTGCTGTGTT 5260
Qy 248 ---AspGlnProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266

ORIGIN

Db 5261 GAGGATACACCACTAAGGCAACCAAGATATCAAGGCAATTCGCGGTCTGATGAATGC 5320
Qy 267 GluAsnGlyThrGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 5321 GATACGGGTCTCAGAAAGTAAATCTTGCT----- 5353
Qy 287 AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu 303
Db 5354 -----ATCAGTGGGTGCGAGGTAGCAAAATCTCGCAAGTTCTGCTACTCAGTGAAGTTG 5407
Qy 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db 5408 CTTGGTAATACAAACCAAGTGTGTGTAGTAACATCAGAAAA----- 5449
Qy 324 GluGlnValValLeuLysArgSerLysSerLysArgLysThrAspLysLysMet 343
Db 5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGTTCGAAAA-----AGAAAGTTGTTA 5503
Qy 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp----- 360
Db 5504 CTTGAAAAAATATGATCGCCGGATATTGAGTACAAATGGTGGCAACCTCTGAAAAATGCT 5563
Qy 361 AlaLysLeuCys----- 364
Db 5564 TCCAAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623
Qy 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db 5624 ACTCCATTTAAGGTAAGCAGAGAAACAGAGAAATTCAGGTGTTGACAGATTGTACCA 5683
Qy 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db 5684 TCATCTCTTGTAAGTCTACAA-----GAAGGTATCAAGGATGATGATGATGATGATGAT 5737
Qy 400 CysGluAspAspArgSerThr-----IlePro 408
Db 5738 AGTAAG-----AGATCAACTCTCGCACTCTTTTATCTACTGGAACAGATCTGTCTCT 5791
Qy 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db 5792 TGTCTCT-----CCGGTACTCAGAGACA-----GAGAGGAAG 5824
Qy 429 LeuLysSerSerLysAsnLysThrLysArgLysTySerAspValValAspAspGlySer 448
Db 5825 CTCAGTTTATCCCAAGAGAGACAAAGAA-----CCTGTAATCGATATATGGGAAG 5875
Qy 449 SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly 462
Db 5876 AGCACTGTATCAGTTTATGATACGGCATTTGATGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5935
Qy 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db 5936 CTTCCATGAACACAGATATCCAAATCCGAGACTTATTGAATGGGAAAGGGTGGCGGT 5995
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp----- 499
Db 5996 TTATTTCACAAACCGTTTGGCTTCAGATGGATATTTTCAGAAAAATATCTCTCTCAGTTAAT 6055
Qy 500 -----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db 6056 GATAAGCCGATAACATCTTTCATTTCCAGACAAATGATATGATGAGTCAAGAGACGCG 6115
Qy 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu 528
Db 6116 GAACCAAACTCTCTTCGAGATTTTATGTTCTCTTCTAAATCCAGCTCAGTGGATGTTG 6175
Qy 529 SerLysGly-----LysThr 533
Db 6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAAACCAACACATACAAACAGATCGTCT 6235
Qy 534 HisSerAlaAlaSerThrLysTyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db 6236 TTCTCGAACTTGAAGTAAAGATACCCCTCTCTTCTACTGAAGTGGCGATTTATCTCGG 6295

Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db 6296 GTGCTGCAAAAGGTACAAGCTTTCATCTATTTTCTCAATCGCCATGCTTCTCATTC 6355
Qy 573 K----- 573
Db 6356 TGTTTTCCAACTTGCACTATTTCTTGATATATCCAAACTGGTGTGCTTCACTGATTGGAT 6415
Qy 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
Db 6416 CAGCATGATACCGGATTTGATTTACTTTGGCAGGATGCTTCTGTCGCGATAGAAAGGGA 6475
Qy 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysLysLysLeuG 604
Db 6476 AGACTGTTATGTCGTCAGAACATCATGAGCAGCAGCAGAGCCAAAGTCCAGTATGAAAGG 6535
Qy 604 LuValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeu 624
Db 6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592
Qy 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 6593 TGGCCAAAAACAGTACGAGAGGTGCTTCCCGCAAGAGAGAGATGTTAGCAACAAC 6652
Qy 643 snArgIleGlnSerLysThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db 6653 AGCCATCAACAAGAAACAGCACAAATCCAAAGATGCTCTACTGATTGATCTCAATGAA 6712
Qy 663 LysSerAspTyAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA 681
Db 6713 CCTAGATTAACGGGATCTCATCTTGGAGACAACACACATCAAGACCACCAAAACCGTGA 6772
Qy 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 6773 GTAGCAACGAGGAGGAGAGAA-----CATTTCTTATGGGAAACAGACAGCAAACT 6823
Qy 701 roHis-----ProGlnAsn-----PheGlns 708
Db 6824 CTCATGACTTCTCCCAATAAGTCAGCCTTATGTCCTCTCCGTTGGGATCTTCTCTC 6883
Qy 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaLysS 728
Db 6884 CTACCCCAAGAAACCGA-----GCCAGCT 6907
Qy 728 erProLeuPheSerHisHisAspAspGlnTyIleAlaGluAlaProThrGluHisTrp 748
Db 6908 CCATCCGGTTTTCTGCTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG-----G 6961
Qy 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 6962 GTAATCAGAACCTTCTCCATCTCTCATTTCCGGTATTAGTGCT----- 7005
Qy 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Db 7006 --TGTGATACTTGCAGAGTGTTCCTAATCAATACAGA----- 7041
Qy 785 euThrSerThrHisValMetGlySerSerSerSerSerSerSerSerSerSerSerSer 805
Db 7042 --GAAGCTTCTCATCCCAATTTGCCATCTTCC-----ATGATAC 7078
Qy 805 laProLeuAspArgTyAlaGluArgAlaValAlaAsnGlnValHisAlaArgAsnPheProS 825
Db 7079 CACCACAGAGTCAATACAGCCAGCTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135
Qy 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgAsnAlaGlyGlnValV 845
Db 7136 GTAAGCTTTTCA-----CAGGCATCTAAACATGAATATACATGGAACCTCACTTTGTTG 7189
Qy 845 allLeuTyProLysGlu----- 850
Db 7190 CTGCCAACGGAGCAAAATATGGCCCTAATCCAGAAATTTTCAATTTGCTGCTCAAAACATG 7249

```

QY 851 -----SerM 852
Db 7250 CTCGCGGGTTAGTAGTAGTAGAGCCAAATAGTAACCTTTCTAGTAGGAGCTCTA 7309
QY 852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPro----- 870
Db 7310 TACCGGCAATGCATCTACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 7369
QY 871 --AsnTyrGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 7370 ACCACACGGAACACTAAATTTACTAAAGACATTTTCGCCGACCAACCACTCCANAG 7429
QY 888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSer 908
Db 7430 AGTTTATAGAGCTTCAACACGGGAGCTCTAGTAAGTCAGCTAGCTAGCTAGCTAGCTAG 7482
QY 908 lylsileProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db 7483 --CAGATACCT-----TTTGATCTATACAGCAGAGATTCACACAGAGCCCTCCGGA 7534
QY 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluLeuA 947
Db 7535 AGAGTTTCCCATCCTCCATCCACTATTGGGACG-----TCTTCACTTTCATTTCAAAA 7588
QY 947 laAsnTyrSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db 7589 CTTCTAGGAGTCTCAT-----TTTGATCTATACAGCAGAGATTCACACAGAGCCCTCCGGA 7605
QY 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 7606 -----CATCAAGAGAAGAACCAACCAAGAGAAAGACACC-----TTTG 7642
QY 987 erAlaLysTyrAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 7643 CTCCTGTTTCAATCTACTCATGAAAGCGGTGTTCGACGACGACCAACGACCAAGT 7702
QY 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTyrThrArgGlyLysGlyLysM 1022
Db 7703 TCCAGCTGCTGGGAGCTCATCTCC-----AspArgPheValArgGlnAsp-----A 7729
QY 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 7730 TGATGCTCTCTTCAAAATTTTCAATGACGATGACGATGACGATGACGATGACGATGACG 7789
QY 1034 -----TTTGATCTATACAGCAGAGATTCACACAGAGCCCTCCGGA 7908
Db 7790 AGAGCTGCAATACATGCTCTGCGGACCTGTGAGAAACAGCTTCTGACCACTTGTGT 7849
QY 1035 ysileThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db 7850 GCAGCGTCAATAGAACCTGCTGATTTTCACTTCCATCTCTGAACTGGAATGTTTACATG 7908

RESULT 9
AR368884 17341 bp DNA linear PAT 12-SBP-2003
DEFINITION Sequence 1 from patent US 6376751.
AUTHORS Sung,Z.R., Aubert,D. and Chen,L.
ACCESSION AR368884
VERSION AR368884.1 GI:34603268
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 17341)
AUTHORS Sung,Z.R., Aubert,D. and Chen,L.
TITLE Nucleic acids encoding EMF1 that control reproductive development in plants
JOURNAL Patent: US 6376751-A 1 23-APR-2002;
FEATURES Location/Qualifiers
1..17341
/mol_type="genomic DNA"
ORIGIN

```

```

Alignment Scores:
Pred. No.: 9,7e-08 Length: 17341
Score: 298.00 Matches: 225
Percent Similarity: 35.53% Conservative: 152
Best Local Similarity: 21.21% Mismatches: 360
Query Match: 324
DB: 6 Gaps: 50

US-09-721-114-2 (1-1057) x AR368884 (1-17341)
QY 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db 5090 TCTAATGTTAGGAGCAGAGCAGAAAGCTGTGAGTCCGGACAG-----GTAGGA 5140
QY 218 AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu 234
Db 5141 AACCAACAGATCTTAAGAAACAACTAAACCAACCATCGATGATATTAGCAGCTCGAAGAG 5200
QY 235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla----- 247
Db 5201 AAACAAATGTGGATCAGCTGTGACACAGTTCCGGCTCATCTGAAATCTCTGGTGGTGT 5260
QY 248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db 5261 GAGGATACACCACTTAAGCAACCAAGATCATAAAGCATTCGGGTCTGATGGAATGC 5320
QY 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 5321 GATAACGGGTCTATCAGAAAGTATAAATCTTGTCT----- 5353
QY 287 AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu 303
Db 5354 -----ATGAGTGGGTGCGAGGTAGGAATCTCGCAAGTTTCGTCTACTCAGTGAGTTG 5407
QY 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db 5408 CTTGTTAATACAAAACCAACAGTGTGTGTAGTAATCATACAGAAAA----- 5449
QY 324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db 5450 GAAGAGTCTGCTTTGAAGAGGAATCATGTTAGAGGTCTGAAAA-----AGAAGTTGTTA 5503
QY 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp----- 360
Db 5504 CCTGAAACAAATATTGTCTAGCGCGATATTGACTACATATGGGTGCAACCTCTGAAATGCT 5563
QY 361 AlaLysLeuCys----- 364
Db 5564 TCCAAAGTTGTGCTCTGATCAAGGTATATAGTAATCACTGATAGTGGGTTCACAGA 5623
QY 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluIleLeuAsn 379
Db 5624 ACTCCATTTAAGGTTAAGCAGAGAACAGAGATTTGAGTGTGTGACGAGTTGTACCA 5683
QY 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db 5684 TCACCTTCTCTGTGAAACCTTCACAA-----GAAGGTATCATAGCAGCATGATCATCT 5737
QY 400 CysGluAspAspArgSerThr----- 408
Db 5738 AGTAAG-----AGATCAACTCTCTCGGCACTCTTTATTCACCTGGAACAGATTTCT 5791
QY 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db 5792 TGTCTCT-----CCGGTACTCTCAGAGAAC-----GAGAGGAAG 5824
QY 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db 5825 CTCAGTTTACCCAGAGAAGAACCAAGAG-----CCTGTATCATGATATGGAAG 5875
QY 449 SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly 462
Db 5876 AGCAGCTGTGATCAGTTTATTAGTAACGGCAATTGATGCAAGTCAAGTTAACTCGCATACTG 5935

```

QY 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db 5936 CCTTCATGACACCATGATCCCAAACTCGAGACTTATTGAATGGGAAAGGGTGGCGGT 5995
QY 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp 499
Db 5996 TTATTGACAAACGGTTGGCTTCAGATGGATATTTTCAGAAAATATCTCTCAGGTTAAT 6055
QY 500 -----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluLysLeuSer 515
Db 6056 GATAAGCCGATAACATCTTTGCAATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGG 6115
QY 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly-----Leu 528
Db 6116 GAACCAAACTGTCTCGAGATTTAGTCTCTCTTAATCCAGCTCAGGTGGATGGTG 6175
QY 529 SerLysGly-----LysThr 533
Db 6176 AGAACTGGAGTAGATATTGTGACTTCAGAAACAACAACCAATACAAACAGATCGTCT 6235
QY 534 HisSerAlaAlaSerThrLysThrGlyGlySerThrArgAsnGlyGlnAsnIleHis 553
Db 6236 TTCTCGAACTTGAGCTAGATACCCCTCTCTTACTGAAGTTGCGGATTTATCTCG 6295
QY 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHis 573
Db 6296 GTGCTGCAAAAGGTACAAAGCTTTGCATCTATTTTCTTCAATCGCATGCTTTCTCATTC 6355
QY 573 r----- 573
Db 6356 TGTTTTCAACTTGCACCTATTCTGTATATTCAAACTGGTGGTCTTCAGTATGGAT 6415
QY 574 -----Ala-LysValSerProAlaGluHisAspIleG 594
Db 6416 CAGGATGATAACGGATTGATTACTTGGCAGGATGCTTCTGGTGAGATAGAAAGGGA 6475
QY 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG 604
Db 6476 AGACTGTTATGGTCCAAAGACATCATGAGCACCAAGAGCCAAAGTCCAGATAGAAAGG 6535
QY 604 InValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 6536 AGACTACGACTGAAGACCAAAACAAC--CATGATATTCATGGAGATAGTGGAGCTCA 6592
QY 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 6593 TGGCCAAAACCAAGTACGAGAGGTGCTTCCCGACAAAGAGAGATGTTAGCAACAAAC 6652
QY 643 snArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspG 663
Db 6653 AGCCATCACAAGAACACAGCACAAATCCAGATGCTCTACTGATTGATCTCAATGAA 6712
QY 663 LysSerAspTyrAlaSerSerValPheAspThrAsnSerGlnIn-----LysSerLeuA 681
Db 6713 CCTACGATAACGGGATCTCACTTGAGGACAAACACATCAAGCACCAACCAACCGTGTA 6772
QY 681 LysGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 6773 GTAGCAACGACAGAGGGAAGAA-----CATTTCTCTATGGGAGACAGCAACT 6823
QY 701 roHis-----ProGlnAsn-----PheGlnS 708
Db 6824 CTCATGACTTCTTCCCAATAAGTCAGCTTATGTCCTTCTCGGTTTGGGATCTTCTCTC 6883
QY 708 erThrGlnGlnGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db 6884 CTACCCCAAGAAAACCGA-----GCCAGCT 6907
QY 728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaproThrGluHisTrpG 748
Db 6908 CCATCCGGTTTCTGGTCAACAACCTGTGAGGGCTTGGGAATTTGCCAACACTGTG-----G 6961

QY 748 lYArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 5962 GTAATCAGAACCTTCTCCATCTCTCATTTGGGTATGAGTGCT----- 7005
QY 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Db 7006 --TGTGATACTTGCAGAGTGTCTCTAATCAATACAGA----- 7043
QY 785 euThrSerThrHisValMetGlySerSerAsnTyrAlaSerArgGlnProValIleA 805
Db 7042 --GAAGCTTCTCATCCATTTGGCCATCTTC-----ATGATAC 7078
QY 805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db 7079 CACCAAGAGTCAATACAAAGCCAGTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135
QY 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnVal 845
Db 7136 GTACGCTTTC-----CAGGCATCTAATGAANAATACATGGAACCTCACTTTGTTG 7189
QY 845 aLeuTy-ProLysGlu----- 850
Db 7190 CTGCCAAACGGGAAGCAAAAATGTGGCCCTAATCCAGAAATTTTCATTTGGCTGCAAAACATG 7249
QY 851 -----SerM 852
Db 7250 CTGCTGGGTTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 7309
QY 852 esProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro--- 870
Db 7310 TACCGGCATGTCATCTACTCAGCCTTCTGGATCTCTGCTGAGGTCAACGACTCCGCTG 7369
QY 871 --AsnTyGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 7370 ACCAAACAGCGAAACACTAAATTTACTAAAGACATTTTCCGCGACCAACAGTCCCAAG 7429
QY 888 InTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSer 908
Db 7430 AGTTTATAGAGCTTCAAACAGCGGACTTAGTAAGTCACTCACTCACTCACTCACTCACT 7482
QY 908 lYLeIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db 7483 --CAGATACCT-----TTGATCTTATACAGCAGAGATTCACACAGAGCCTTCCCGA 7534
QY 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db 7535 AGAGTTTCCCATCATCTCCACTATTTGGAGC-----TCTTCACITTTCAATTTCAAATG 7588
QY 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db 7589 CTTTCATGAGTCTCTCAT----- 7605
QY 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 7606 -----CATCAGAGAGAAACCAAGAGAAAGACACC-----TTTG 7642
QY 987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 7643 CTCCTGTTTACATACTCATGAAAGCCGGTGTGTTGCAAGCAGCAACACCAAGCGAAGT 7702
QY 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db 7703 TCCAGCTGCTGGAGCATCGAATTC-----A 7729
QY 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 7730 TGATGCTTCTTTGAATTTTCATGACGGATAAGAAAAGAAACAAAGAGAAAGCAG 7789
QY 1034 -----IleC 1035
Db 7790 AGAGCTGCAATACAAATGCTCTGCGGACCTCTGAGAACACAGTTCTGAGCCCATTTGTG 7849
QY 1035 ysIleThrAsnLysAsnProAlaAspPheThrIleSerAsnAspAsnGluTyrMet 1054

```

Db      7850 GCACGTCATAGAACCCCTGCTGATTCACCATCTCTGACCTGGGAATGTTTACATG 7908
|||||
RESULT 10
LOCUS   AT15N18      101665 bp      DNA      linear      PLN 13-APR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone F15N18 (BSSA
project).
ACCESSION AL163815
VERSION   AL163815.1 GI:7573398
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 101665)
AUTHORS   Bevan,M., Hilbert,H., Braun,M., Holzer,E., Rudd,S., Lemcke,K. and
Duesterhoeft,A., Bancroft,I., Mewes,H.W.,
Unpublished
JOURNAL   2 (bases 1 to 101665)
REFERENCE Direct Submission
AUTHORS   Submitted (13-APR-2000) MIPS, at the Max-Planck-Institut fuer
TITLE     Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
JOURNAL   lemccke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT   Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source            1..101665
                     /organism="Arabidopsis thaliana"
                     /mol_type="Genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
     gene              6461..8797
                     /gene="F15N18_10"
     CDS               join(6461..6536,7295..7792,8271..8797)
                     /gene="F15N18_10"
                     /note="similarity to predicted proteins in castor bean,
                     Arabidopsis thaliana and alfalfa."
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CA887702.1"
                     /db_xref="GI:7573399"
                     /db_xref="SPTREMBL:O9LYE7"
                     /translation="MKGSLSLFLVLIATITTSVCFSDGMLPNEDFELGPKPSDMKG
                     TVINKKAIPSWELGPFVEYIKSGQGMLLVVPACKFAIRLGNESAIKQRLNVTG
                     MYSLTPSAARTCAQERLNISVAPDSGVIPIOTVYSSGMDIYANAFQAESNVAIV
                     IHNPGREDACGPLIDGVAIKALYPPRTNKNILKNGSGPEEGPYVLNATGVLVPP
                     FIDDSHSPANWVESLKAIKYVDVEHFSVPQGRRAVELVAGKESAIQAQVATVUGKT
                     YVLSFVGDANACQGSVVEAFAGKDTLKVYPYESKGGKGFKRASLRFVAVSTRTRVM
                     FYSTFYMSRSDFFSSCGPVIDDVKKLSARKP"
     exon              6461..6536
                     /gene="F15N18_10"
     intron            6537..7294
                     /gene="F15N18_10"
     exon              7295..7792
                     /gene="F15N18_10"
     intron            7793..8270
                     /gene="F15N18_10"
     exon              8271..8797
                     /gene="F15N18_10"
                     /number=3

```

```

10281..14062
/gene="F15N18_20"
/join(10281..11965,12429..12549,13145..13204,13313..14062)
/gene="F15N18_20"
/note="similarity to putative transcription elongation
factor S-II, Arabidopsis thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CA887703.1"
/db_xref="GI:7573400"
/db_xref="SPTREMBL:O9LYE6"
/translation="MAEINGSMQLVGMKSLPQTTLGGSGASEAPNKKVQRPMLQQLSP
ASNGLIHPTKILSOFTLHSLMHGKATQESAPOKPAKPVVKKQHVPPPORSVKAM
EHNVSSEKMEESLASALVKDDSDSPKENIGTVETPVITQENTOSFOPASPAS
ISVPGQGTMEPTSVSSVOKDSEIPIVDIMMEDVIKENVLKQSDVEFPDRNVPTT
DIIFNDLLHGNLSWLEVDLGETKDYGTGGKSPQPKLASKIEMELYKULPGG
VANKYRKERSLLFNLDKNNPRLRVMSEISAEKLSMTABELSKELSSQWROAK
AEEMAKVVLQTDIDVRSLVKTHKGFVEIIPVDRGTVDVGGIIMRSKRPRAK
SHSVTKALDEAAKADNKSSTPSTBEIDPMQGLIGDDELKDFEFLPPIVSLDBFM
BSLSDPEPPSPHGNSSQMSPEKSDSEAGSDSKSPKGSPEKLSLPEAKPEKID
EYTPFDANVKVDDDISRVEKAAALSDDKGERAWDGILOLSMSVVPVAGIFKSGEKA
ETSENPAMVEYKGRVLSGPKFIQELPKSRTRALMVYLAYEDGISQSGSLIEVI
DSYVADQRVGABPASGVELYLCPTGTGTLDLINKVISOQLDEKSLDGLGVGVVW
RSAYVPKSGSGKRSFSSSISGKTVLPVKKQRHVTEKPLVAGNRNHHGKLIHK
HDTAADDDVDPGPGVARSDEDDLPFNSSVVPSSPOPLPAQSKSLDQVKRLIHK
YKGSASTYDDDDDDDDIIPWQPHVPESHQLPPPPPPPLGFRPEVFPAPPDQGYDNQNGG
SSOHYERNQSRNRGP"
10281..11965
/gene="F15N18_20"
/number=1
11966..12428
/gene="F15N18_20"
/number=1
12429..12549
/gene="F15N18_20"
/number=2
12550..13144
/gene="F15N18_20"
/number=2
13145..13204
/gene="F15N18_20"
/number=3
13205..13312
/gene="F15N18_20"
/number=3
13313..14062
/gene="F15N18_20"
/number=4
14866..15413
/gene="F15N18_30"
complement(join(14866..15009,15090..15413))
/gene="F15N18_30"
complement(join(14866..15009,15090..15413))
/codon_start=1
/product="putative protein"
/protein_id="CA887704.1"
/db_xref="GI:7573401"
/db_xref="SPTREMBL:O9LYE5"
/translation="MKPGAFALNPHAAASVVPISKRVYGGDDGLVFAAKSPTVEYSM
PKKSEMAKQIRDDDLDEMIDMDIEVLLVTFPSLQSEITDVLNANGGDLKLEATIE
MLNQLEIYSTESENLPELIDIGDISPSGPKSKSTVAASTSSVIPNAPVSA"
complement(14866..15009)
/gene="F15N18_30"
/number=1
complement(15010..15089)
/gene="F15N18_30"
/number=1
complement(15090..15413)
/gene="F15N18_30"
/number=2
16227..18163
/gene="F15N18_40"

```


Qy 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu 528
Db 60821 GAACCAACGTGCTTGGAGATTTTGTCTCTTCTAAATCCAGCTCAGGTGGATGGTGG 60880
Qy 529 SerLysGly-----LysThr 533
Db 60881 AGAAGCTGAGTATGTTGACTTCAGAAACAAACACCAATACAAACAGATCGTCT 60940
Qy 534 HisSerAlaAlaSerThrLysThrGlyGlyLysThrArgAsnGlyGlnAsnHis 553
Db 60941 TTCTCGAAGCTGAAGCTAAGATACCCCTTCTTCTACTAGTTGGGATTTATCTCGG 61000
Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHis 573
Db 61001 GTGCTGAAAAGGTACAAAGCTTTGCACTATTCTTCTCAATGCCCATGCTTCTCATTC 61060
Qy 573 r----- 573
Db 61061 TGTTTTCCAACTTCACCTATTCTGTGTATATCCAAAGCTGGTGTCTTCAGTATGGAT 61120
Qy 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
Db 61121 CAGGATGATTAACGGGATTTGATTTACTTTGGCAGGATGCTTCTGTGTCAGATAGAAAGGGA 61180
Qy 584 InitLeuMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db 61181 AGACTGTTATGGTCCAAAGAACATCATGAGCAGCACCAGAACCAAGTCACGATAGAAAG 61240
Qy 604 luValThrArgGluLysGlnThrMetIleAspIleProMetAspIleValGluLeuL 624
Db 61241 AGACTAGGACTGAAGAGCAAAACAAAC---GATGATATTCCAAATGGAGATAGTGAGCTCA 61297
Qy 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 61298 TGGCCAAAACACAGTACGAGAGGTGCTTCCCGCAAGAAAGAGATGTTAGCAACAAAC 61357
Qy 643 snArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspG 663
Db 61358 AGCCATCACAAGAACACGACACAAATCCAGAAATGCTCTACTGATGATCTCAATGAAA 61417
Qy 663 LysSerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA 681
Db 61418 CCTACGATAAGGGATCTCACTTGGAGACAAACACATCAAGACCAACCAACCGGTGA 61477
Qy 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 61478 GTACCAAGCAGGAGGGAGAA-----CATTTTCTATGGGAGACAGCAGACT 61528
Qy 701 roHis-----ProGlnAsn-----PheGlns 708
Db 61529 CTCATGACTTCTCCCAATAAGTCAGCCTTATGTCCTTCTCCGTTTGGGATCTTCTC 61588
Qy 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db 61589 CTACCCCAAGAAACCGA-----GCCAGCT 61612
Qy 728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db 61613 CCATCCGGTTTCTGGTCACAACTGTCTGAGTGGTGGGAATTTGCCAACTGTG-----G 61666
Qy 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 61667 GTAATCAGAACCCCTTCCCATCTCTCATTTCCGGTATTACGTGCT----- 61710
Qy 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Db 61711 --TGATGACTTGGCAGAGTGTCTTAATCAATACAG----- 61746
Qy 785 euThrSerThrHisValMetGlySerSerAsnTyrAlaSerArgGlnProValIleA 805
Db 61747 --GAAGCTTCTCATCCAAATTTGGCCATCTTCC-----ATGATAC 61783
Qy 805 laProLeuAspArgTyrAlaGluAlaValAlaAsnGlnValHisAlaArgAsnPheProS 825

Db 61784 CACCACAGAGTCAATACAGCAGCCAGTTTCTTTAAATATTAAATCAGTCAACAAAT---CCGG 61840
Qy 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db 61841 GTACGCTTTCA-----CAGGCATCTAACATGAAATATACATGGAACCTCAACTTTGTTG 61894
Qy 845 allLeuTyrProLysGlu----- 850
Db 61895 GTGCCAACGGGAAGCAAAATGTGGCCCTAATCCAGATTTTTCATTGGCTGCAAAACATG 61954
Qy 851 -----SerM 852
Db 61955 CTGCTGGGTAGTAGTAGTAGTAGTAGTAGCCATAGATAACTTTTCTAGTGAGAGCTCTA 62014
Qy 852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro----- 870
Db 62015 TACGGCATTGTCATCTACTCAGCCCTTCTGGATCTCTGGATCTCAGCTCAGGACTCCCGCTG 62074
Qy 871 --AsnTyrGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 62075 ACCAACCGCAACACATAAATTTACTTAAAGACATTTTCGGCCAGCCACACCTCCAAAG 62134
Qy 888 lntYrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db 62135 AGTTTATAGAGCTTCAACACAGGGGACTCTAGTAAGTCAGCTACTCAACTAAG----- 62187
Qy 908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db 62188 --CAGATACCT-----TTGATCTTATACAGCAAGAGATTCACAAAGACCTTCCCGGA 62239
Qy 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db 62240 AGAGTTTCCCATCACTCCACCTATTGGGAGC-----TCTTCACTTTCAATAAATG 62293
Qy 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db 62294 CTTCTAGGAGTCTCTCA----- 62310
Qy 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 62311 -----CATCAGAGAGAAACCCAGAGAAAGACACC-----TTTG 62347
Qy 987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 62348 CTCTCTGTTTACAATACTCATGAAAGCCGTGTTTTCAGACGACGACCAACGAGGAAGT 62407
Qy 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db 62408 TCCAGCTGCTGGAGCATCGAATTC-----A 62434
Qy 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 62435 TGATGCTTCTTTGAAATTTTCATGACGATAAAGAAAGAAACAAAGAGAGAAAGCAG 62494
Qy 1034 -----IleC 1035
Db 62495 AGAGCTGCAATACATGCTCTGGGGACCTGTGAAGAACAGTCTTGAGCCCAITGTGT 62554
Qy 1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db 62555 GCAGCGCTCATGAAACCCCTGCTGATTTCCATTCCTCACTGACCTGGGATGTTTACATG 62613

RESULT 11
MUSCVCOLORA
LOCUS 7036 bp mRNA linear ROD 20-OCT-1999
DEFINITION Mus musculus cyclophilin-related protein mRNA, complete cds.
ACCESSION L04289
VERSION L04289.2 GI:6073857
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 7036)
 Anderson, S.K., Gallinger, S., Roder, J., Frey, J., Young, H.A. and
 Ortado, J.R.
 A cyclophilin-related protein involved in the function of natural
 killer cells
 Proc. Natl. Acad. Sci. U.S.A. 90 (2), 542-546 (1993)
 93133824
 8421688
 2 (bases 1 to 7036)
 Anderson, S.
 Direct Submission
 Submitted (27-APR-1993) Lab of Experimental Immunology, NCI-FCRDC,
 Building 560, Room 31-93, Frederick, MD 21702-1201, USA
 3 (bases 1 to 7036)
 Anderson, S.
 Direct Submission
 Submitted (20-OCT-1999) Lab of Experimental Immunology, NCI-FCRDC,
 Building 560, Room 31-93, Frederick, MD 21702-1201, USA
 Sequence update by submitter
 On Oct 20, 1999 this sequence version replaced gi:192866.
 Location/Qualifiers
 1..7036
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 /cell_type="NK"
 /tissue_type="blood"
 /dev_stage="adult"
 63...4511
 /notes="peptidyl-prolyl cis-trans isomerase; transmembrane
 protein"
 /codon_start=1
 /product="cyclophilin-related protein"
 /protein_id="AA37500.2"
 /db_xref="GI:6073858"
 /translation="MVLGGSGRAFGSNAVPRRRPAALAMCAQDRPQCHFDIEI
 NRPVGIMFQPSDICPTKCNFLCLSGEGLGKTKKLCYKSTPHRVVKNFMI
 QGDPGEKGESYGYFKDENFLKDRAPFLLSMAHRGKHTNGSQFFITTPAP
 HLDGVHVLGVIGPEVIOENLXATDAASPVADRVVIDGVLATKLTKDVFERR
 KPTCEGSDSSSSSSSESSESEVERETRRRHRRPKVRHAKKRKEMSSSE
 PRKRTVSEGVSESDVNEKSVDSNTREKPVTPPERIPVPENRPLLRDMPAIT
 VEPEQNPVAPVDSQKESVSGKRIKGRGTIRVHTPPRSRSHSEKDDSSSTPP
 HKKEEMORLAVNPPSGEKGSKGLKSLDPCSRWDERLSLQSRKSWNGYISDLSTA
 RSDGHHKHKRKKPKKKKAKQKCHRRQTKKRIWMPDLPSRSPTRHMKSSC
 VERRRASSSSHSKSDWSKDDGASSTHSSRDSYRSKSHSRSDSGSSRSRA
 VSKSSSLNRSRSGSPRTSISPKPAQISENPKVTEPLRPSPQNGVLI
 QPVAENI PVLPLSDPPSRWPKGOKPKPKSVERTIOENKAKTTHLLPVOSTYSLNI
 KATVSSSHHKKRKPESDGSAYSKYSDSSGSGSGSKSSRSRSTRSRSSRS
 LPTSRSLSPSRSPSPKISDGOHRSRSTSVSSDDGRAMPERSNRKSVTSRK
 RHRSNKKGLRGGEPKPEWECPRSKENSEDHSDDSVSKNGCAGKWSSENSE
 VEALNDQKGRGEGKPEWECPRSKENSEDHSDDSVSKNGCAGKWSSENSE
 QVYTKSKGDPGRSGKEASDSESEVQSHIKAPPAKPTSTPLPGSDGAWKS
 RPQSSASRSESSCNLGNTRGPQKESKODLKGDTKRAKESKAKKDKKAPK
 RKOAFHWQPLFGDDPEEMNGKQVTPDKEKRVHSEKAEVQDGI PNVEKTCDEGS
 SPSPKPKLQDPLAEGGHPSSCAPLKVDENTASPPSAQHLEHGGGGEVLQ
 TDNMETCTDRTSPAKGEVYSLANRLDSEPVNLIPEQDCMAHPRAGGQESMS
 ESKTLGESGVQDSSTSVTSVPVTSKKGAEKSKQNLIDKPKPLQGVGNLVSATTT
 SSALDKALSTVEVPKQGLRIEIKSNKVRPGSLDFEVRKTLARLNRRPNRSESSD
 QTPSRDQTSRSPSRSPSKSTSRHRTSRSSYSRSHSRSSSTYSRSTRSRSR
 RQWYSGRTSRSSYSRSHRTSRSSYSRSHSRSSYSRSHSRSSSTYSRSTRSR
 SQRSDSVHRGSRNRSRSGRSGSDSRSSDRSYSHRSPSESRYS"

source

81 IlePheHisaspGlnLysLysCys-----AspGluHisLys-AlaSerSe 88
 515 GTTTTTCATAACTACAAAGCTGCTCCACACCTGATGGGTTTCATGTTGTTTGGACT 574
 89 -----
 575 GGTAAATATCTGTTTGAAGTAATTGAACAGATTGAAAACTGAAACAGATGCTGCAAG 634
 95 rSerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLysLysLe 115
 635 CAGACCTTAT-----GCAGATGCCGAGTTATGACTGTGGGTGCTGGCCACAAAGTT 688
 115 uLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSe 135
 689 G---ACAAAGATGTTTGTGAGAAAAAGAGAAACCAACCTGCTTCAGAAAGC----- 740
 135 rAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGl 155
 741 -----TCGGACTCTTCTTC 754
 155 nLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeu-- 174
 755 CGTCTCTCTTCTCTCTCAGAGTCTCTCTCAGAGAGTGAAGTTCAGAGAGACAAATCAG 814
 175 -----ProLys----- 176
 815 AAGCAGAGACATAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 874
 177 ---SerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAl 195
 875 GAGCAGTTTCAGAGAAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 934
 195 aAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAs 215
 935 GAGTGATGTGAAT-----GAAAAAGATC 958
 215 pValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluVa 235
 959 AGTTCACTCAACACTAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAG 1018
 235 lProGln-----lleThrTrpHisIleGluVa 244
 1019 TCCCGAGAACCGATTTTACTTAGAAGAGATATGCTCTCTCTCTCTCTCTCTCTCTCT 1078
 244 lAsn-----GlyAlaAspGlnProProSer----- 252
 1079 GAACATTCAGATGTTGCACCTGTTTGAAGTATCAGAAACCTCTCTCTCTCTCTCTCT 1138
 253 -----ThrProLysLeuSerGluValVa 260
 1139 ACGGAAATCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198
 260 lleuLysArgAsnGluAspGluAsnGlyThrGluGlnThrLeuValAlaGluGlnCy 280
 1199 CTCTGAGTCCAAAGATGATGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1258
 280 sAsnLeuThr-----LysAspProAsnProMetSerGlyLysGluArgAspGlnVal 298
 1259 GAGACTGAGAGCTACAGGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1318
 298 aGluGlnCysAsnLeuThrLysAspProValSerGlyGlnLysCysGluGlnI 318
 1319 TGACCCCTGTTTCAGCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1378
 318 eCysAsnGluProCysGluGluValValLeuLysArgSerSer----- 332
 1379 CTATATGATGATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1438
 333 -----LysSerLysArgLysThrAspLysLeuMetLysLysGlnHis----- 348

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 7,95e-05 Length: 7036
 Score: 241.50 Matches: 266
 Percent Similarity: 33.26% Conservative: 162
 Best Local Similarity: 20.67% Mismatches: 428
 Query Match: 4.37% Indels: 432
 DB: 10 Gaps: 60

QY 227 iAspValGlyAlaLeuProGluVal-----ProGlnIleThrTriHisIrl 242
Db 799 AGACTCAGACACACACACAGCAATAAACAACCTCAAGGTTTAAACCTGCGCCGCAAGT 858
QY 242 eGluValAsnGlyAlaAspGln-----ProProSerThrProLysLeuSerGluVa 259
Db 859 CTCACTGACACCTTCTCCAGAGATATGTCACACAGTGTCACACCCATCACTCACTCT 918
QY 259 lValLeuLysArg-----AsnGluAspGluAsnGlyLysThrGluGluThrLeuVa 276
Db 919 TACTCAGACACACACACTGGAAGCAGCAGCAGCAGCAATCTCAAGTCAACATAGT 978
QY 276 lAlaGluGlnCysAsnLeu---ThrLysAspProAsnProMetSerGlyLysGluArgAs 295
Db 979 CACATCAACTCTTCAACTCTTCAAGTGAAGTCACTCACTCTCT-----GCACGACAAATGTC 1032
QY 295 pGlnValAlaGluGlnCysAsnLeuThrLysAspProLysPro---ValSerGlyGlnly 314
Db 1033 TCAAGTGACATCATCTTTCAGAT-----AAAAGAACCAAGCCATCTACTTCTGGTCTCAG 1086
QY 314 sCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSe 334
Db 1087 CAGTACATCACTACTACTACACAGAA-----GTGCTAACCCAAACATCTCTCTAC 1134
QY 334 rLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAl 354
Db 1135 AGACTCTGCTCCAGAAACACAACTCTGAGGATTACTCAGAACTCTACACACATACCAC 1194
QY 354 aGlnAlaAspValSerAspAlaLysLeuCysArgLysProLysLysValArgLeuLe 374
Db 1195 GAAAGTCTCAACACATCTACA-----CCCAAAACTGTCCACAGT 1236
QY 374 uSerGluIleAlaAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgG 394
Db 1237 GTCAACACTCATCAATCAAGTCAAAAATGTCAACCTCTCTCCAGAACCAACACACTGA 1296
QY 394 uAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMetGluValSe 414
Db 1297 A---AGCATGACACACAGCAGCAACCTCAACCAACCAACCAACAGAGGTGACCAATC 1353
QY 414 rMetAspIleProValSerAsnHisThrVal----- 424
Db 1354 AACCCCTTCAGCTCTCTCACTTCAACCAATTCAGACAGACAGAAATTCACCAAAACAAT 1413
QY 425 -----GlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysLysSerAs 442
Db 1414 CTCTCTCTGAGACAAACCACTCACAATGCTCCGAACAT-GAGAAGCTCACCCCAAGA 1472
QY 442 pValValAspAspGlySerSerLeuMetAsn-----TriLeuAsnGlyLysLysLysAr 460
Db 1473 CRTCAGATTAATAACCAATGCTCCCAAAAGTACTCTAGTAGACACAAACAAA 1532
QY 460 gThrGlySerVal-----HisHisThr----- 467
Db 1533 CAAGGCAATAACTACTAAAGTTTCTACACCTGACACCAACCACTCTCGATGACACCTT 1592
QY 468 -----ValAlaHisProAlaGlyAsnLeuSerAsnLysLysValTh 481
Db 1593 CATCCAGAAATGCAACACACAGCAGCTCTACTCAGGAAT---AACATCTCTTTACA 1649
QY 481 rProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAs 501
Db 1650 GCCAACATATTCAAGCAAGGGGACAGCAGCAAAATCTCAAAACCAACCAACCAACCAAG 1709
QY 501 nMetHis-LysThr-----AspValCysGlnHisValSerGluIleS 515
Db 1710 TGAACATCACTCACTCTCTCAGCTCTCTCACTGAGCAAAATTCAG-----ACAGAGACAA 1763
QY 515 erThrGlnArgCysSerSerLysGly---LysThrAlaGlyLeuSerLysGlyLysThrH 534
Db 1764 GTTCCCAAGAACAAACTCTCTCTGGGAGAAAGGACATCATGTGTCACCATGAGCA 1823

QY 534 iSerAlaAlaSerThrLysTyrrGlyGlyGlySerThrArgAsnGlnGlnAsnIleHisV 554
Db 1824 GCTCAGCTCCAGGTACACACACATGTTATCCACCAATCTCCCAATCAAAGTACTCTCAG 1883
QY 554 al-----LeuSerAlaGluAspGlnCysGlnMetGluThrGluA 567
Db 1884 TAGACACAGGCAACACACATCAGTAACAGCTCAAGTTGTACACCTCCCTCATCACAAGA 1943
QY 567 snSerValLeuSerHisSerAlaLysValSerProAlaGluHisAspIle-----G 584
Db 1944 CCTCATTCAGACCTTCATCCAGATATGTCAACAGTGTCAACACCCATCACTCAACTC 2003
QY 584 lNileMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db 2004 AAATTTTGTCA-----ACCCTTCTCAGAGCCCAACACACTGGAACATGG 2048
QY 604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 2049 GGACAGCAGCAACCTCAAACTACACAGCCAGTGT-----GTAAACCATCAAA 2099
QY 624 euAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnA 644
Db 2100 CTCCTTCAGGCACCTTCAGGTGACCAAAATTCAGACAGACCAAGT----- 2143
QY 644 rGleGlnSerLysThrThrAlaAspAspCysValIleValAlaLysAspGlyS 664
Db 2144 -----TCCCAAGAACATCTCTCCGCGACAGACCACTTCACATGCTCCCAATAA 2198
QY 664 erAspTyrrAlaSerSer-----ValPheAspThrAsnSerGlnGlnLysSerLeuA 681
Db 2199 ACAGCTCAGCTCCAAGTACACACACATGTTATCCACACATCTCTCCCTCAAGTACT 2258
QY 681 laSerGlnSer-----ThrGlnLysGluGlnGlyHisLeuAlaLeuThrThrG 698
Db 2259 CAGGAGACACACACACATACACGAGGAGACACTCAAGTCTCACCTGCTCCACACAC 2318
QY 698 lInGluSerProHisProGlnAsn-----PheGlnS 708
Db 2319 AGACCTCAITGACACCTTCATCCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGA 2378
QY 708 erThrGlnGluGlnThrHisLeuArgMetGluGluMetValThr-----I 724
Db 2379 CAGAACAAAGTTCCTCCAAAGAACATCTCTCTGCTGAGACCAACCTCACATGCTCCCA 2438
QY 724 leAlaAlaSerSerProLeuPheSerHisHis-----AspAspGlnTyrrI 739
Db 2439 TCATGAGCAGCTCACTCCCAAGTTCACACACACATGTTATCAACAGCATCTCCACTGAAA 2498
QY 739 leAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluG 759
Db 2499 TTACTCTCAGTAGACACACAGACAC-----ACAAACAGCAATAATGACTCAAGGTT 2546
QY 759 lnPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProG 779
Db 2547 CTACACTCTCCACACAGCTCTCACTCTCATCCAG-----AACATGT 2591
QY 779 lYleGlnAlaValAspLeuThrSerThrHisValMet----- 791
Db 2592 CAACAGTGTGACGACCCCATCATCTCACTCATATTTTGTCAACCCCTTCTCAGAGCCAAC 2651
QY 792 -----GlySerSerSerAsnTyrrAlaSerArgGlnProValIleAlaProL 807
Db 2652 ACCTGGAGAGCAGGGGACCGAGCGAAT----- 2680
QY 807 euAspArgTyrrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrI 827
Db 2681 -----CCTCAAACTA 2690
QY 827 leAlaThrMetGluAlaSerLysLeuCysAspArgAsnAlaGlyGlnValValLeu 847
Db 2691 CCACAAAC----- 2698
QY 847 yrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuA 867

```

Db      2699  -----CCAGTCGGTAAACCA 2711
Qy      867  laSerPheProAsnTyGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnS 887
D5      2712  CATCAACTCTCTCA-----GCCTCTCAGTCACCAATTCAGACAGACAGACAGTCCC 2765
Qy      887  erClnTyralaHisAsnGlnTyLysGlySerThrSerThrSerTyGlySerAsnLeuA 907
D5      2766  AAGAACAAATCTCT-----CCGGCAAGACACACCTCACATGTTCCCAACATGA 2816
Qy      907  snGlyLysIleProLeuThrPheGluAspLeuSer 918
D5      2817  ACAGCTCAGCTCCACAGTACACACACATCTTATCC 2851

RESULT 13
BX842684/c
LOCUS   BX842684          227282 bp    DNA    linear    HTG 03-DEC-2003
DEFINITION  Danio rerio clone DKEY-33122, WORKING DRAFT SEQUENCE, 18 unordered
            pieces.
ACCESSION  BX842684
VERSION    BX842684.1 GI:38677855
KEYWORDS   HTG; HTGS_PHASE1; HTGS DRAFT.
SOURCE     Danio rerio (zebrafish)
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE  1  (bases 1 to 227282)
AUTHORS    Burton,J.
TITLE      Direct Submission
JOURNAL    Submitted (02-DEC-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT    ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            ----- Project Information
            Center project name: zK33122
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 221829 bases at least Q40
            Consensus quality: 223327 bases at least Q30
            Consensus quality: 224335 bases at least Q20
            Insert size: 225582; sum-of-contigs
            Insert size: 226516; 10.5% error; agarose-fp
            Quality coverage: 6.85x in Q20 bases; sum-of-contigs Quality
            coverage: 6.98x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 5980: contig of 5980 bp in length
            * 5981: gap of 100 bp
            * 6081: 11632: contig of 5552 bp in length
            * 11633: 11732: gap of 100 bp
            * 11733: 15379: contig of 3647 bp in length
            * 15380: 15479: gap of 100 bp
            * 15480: 20121: contig of 4642 bp in length
            * 20122: 20221: gap of 100 bp
            * 20222: 44214: contig of 23993 bp in length
            * 44215: 46314: gap of 100 bp
            * 46315: 46914: contig of 2600 bp in length
            * 47014: gap of 100 bp
            * 47015: 50347: contig of 3333 bp in length
            *
            50447: gap of 100 bp
            83115: contig of 32868 bp in length
            83416: gap of 100 bp
            91865: contig of 8450 bp in length
            91866: gap of 100 bp
            97794: contig of 5829 bp in length
            97894: gap of 100 bp
            132361: contig of 34467 bp in length
            132461: gap of 100 bp
            139955: contig of 7534 bp in length
            140095: gap of 100 bp
            144791: contig of 4696 bp in length
            144891: gap of 100 bp
            181114: contig of 36223 bp in length
            181214: gap of 100 bp
            185354: contig of 4140 bp in length
            185454: gap of 100 bp
            188045: 188044: contig of 2590 bp in length
            188144: gap of 100 bp
            191198: contig of 3054 bp in length
            191298: gap of 100 bp
            227282: contig of 35984 bp in length.

FEATURES             source
     1..227282
     /organism="Danio rerio"
     /mol_type="genomic DNA"
     /db_xref="taxon:7955"
     /clone="DKEY-33122"
     /clone_lib="DanioKey"
     1..5980
     /note="assembly fragment:00445
     fragment_chain:1"
     6081..11632
     /note="assembly fragment:00142
     fragment_chain:1"
     11733..15379
     /note="assembly fragment:00061
     fragment_chain:1"
     15480..20121
     /note="assembly fragment:00330
     fragment_chain:1"
     20222..44214
     /note="assembly fragment:00613
     fragment_chain:1"
     44315..46914
     /note="assembly fragment:00042
     fragment_chain:1"
     47015..50347
     /note="assembly fragment:00114
     fragment_chain:1"
     50448..83115
     /note="assembly fragment:01424
     fragment_chain:2"
     83416..91865
     /note="assembly fragment:00380
     fragment_chain:2"
     91966..97794
     /note="assembly fragment:00288
     fragment_chain:2"
     97895..132361
     /note="assembly fragment:01004
     fragment_chain:2"
     132462..139955
     /note="assembly fragment:00527
     fragment_chain:2"
     140096..144791
     /note="assembly fragment:00087
     fragment_chain:3"
     144892..181114
     /note="assembly fragment:02387
     fragment_chain:3"
     181215..185354
     /note="assembly fragment:00172"

```

```

misc_feature      185455..188044
                  /note="assembly_fragment:00206.0"
misc_feature      188145..191198
                  /note="assembly_fragment:00247"
misc_feature      191299..227282
                  /note="assembly_fragment:01877"

ORIGIN
Alignment Scores:
Pred. No.:      0.016      Length:      227282
Score:          234.50     Matches:      213
Percent Similarity: 33.92%  Conservative: 171
Best Local Similarity: 18.82% Mismatches: 439
Query Match:      4.24%    Indels:      310
DB:              2        Gaps:        40

US-09-721-114-2 (1-1057) x BX842684 (1-227282)

Qy      76  CysSerLeuSerArgIlePheHisAspIndinLysCysAspGluHisLysAlaSerSer 95
Db      131791 TGCTGCTGTTCAGGAATATGACATCTCGAAGAACTCTGTACCTCCAGTCAGGAGCTGT 131732

Qy      96  SerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeu 115
Db      131731 AAGACGCATCAGATCCTCAAGCTC-----AATGTGATCCTCAAGAG 131690

Qy      116  LysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGluAsnGlyThrSer 135
Db      131689 AATAGTGTGCAAAAAGGCATCTCTGACAGCAGCATTCGCACAGTGTCCAGTGAGGATCAC 131630

Qy      136  AspGlyCysSer-----IleThrPheValArgSerThrPheValProAlaSerValGI 153
Db      131629 AGAGCATGTAGTGGAAAAGGCAACTTGTGATTCAACAGACCTTC----- 131587

Qy      153  ySerGlnLysValSerProSerThrGlnSer-----SerGlnGI 166
Db      131586 -GGACAGCAAGTAGATACACAAATTTGAACAGTGTACCATCAACCATGCTGCAGTCGAG 131528

Qy      166  yLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCy 186
Db      131527 TGAAGTATCCTCAATATCCCTTAACCAAGAGCACAAAGTACTGTAGATGATAGTGTGA 131468

Qy      186  sAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAs 206
Db      131467 GAACATTCATCATCTGTATAT-----GAAGAGAACAGACACTGATGTAATTCCTCAA 131414

Qy      206  pLeuGlnGlyProAlaGlnAsnTyAspValAlaAlaAsnValSerGluAspAsnThrSe 226
Db      131413 CCTCCTAAAGATCATTTAAGTGTTCAGATCACTCAGCAAGTAGAGAGAGATGAGAGGTC 131354

Qy      226  rValAspValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGI 246
Db      131353 TATATCTGTC-----TTGAGCTCTCTCTCCAAATTTAGCATCATCACTGAAGAGGAGATCA 131300

Qy      246  yAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAs 266
Db      131299 AGATGATGAC-----GTCAACCCCACTATCTCGAGAGCATCCCATGACCAACAGATGA 131246

Qy      266  pGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPr 286
Db      131245 TGAGGAGACCAAGATCCTGAAACA-----AACCAATGTTCTGTGCTTACCCTGTAA 131195

Qy      286  oAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAs 306
Db      131194 ATCTCTCATCTCTCTGAGGCAT-----CTGTC 131168

Qy      306  pProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluVa 326
Db      131167 ACCAAGACCCGAC----- 131152

Qy      326  lValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMetLysLysGI 346
Db      131151 -----AAGGCTTCATCTTCAGGGCGGTCTAAACACTTACAGAAAGTTGCCAGACTCTGA 131099

```

Qy 639 sserAspIleasnArgIleGlnSerLysThrAlaaspAspCysValIleValAl 659
Db 130144 CTCTACATATCCAGAGCATCTCAATTTTCGAGATGAACACAGAC---GAAATCTCTGGA 130088
Qy 659 aAlaLysAspGlySerAspTyrAlaSerValPheAspThrAsnSerGlnGlnLysSe 679
Db 130087 AAAGAGGATGATCAGGAAGAGCATCAAGTGGCAATTCGTGTAAGTCT---TC 130037
Qy 679 rLeuAlaSerGlnSerThrGlnLysGlu----- 688
Db 130036 TGTATCTGACAGATCAAGAAATCTGAGACACACCAATATCAAGATCAGGAGCAAAAT 129977
Qy 689 -----LeuGlnGlyHisLeuAlaLeuThrThrGlnG 699
Db 129976 CAAGATCGGGCTCCAAAGTGTATGCTCGCTCAATCAATCAATCTTCATTGACATCAATAA 129917
Qy 699 uSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuArgMetG 719
Db 129916 GTCC-----AAGATGGAATGAAGAAACAGAAACAGACAGCTTC 129878
Qy 719 uGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspGlnTyrI 739
Db 129877 CAGTGTCTGTCTGTTAAATGGGAAGCTTCACATTGTCA----- 129838
Qy 739 eAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluG 759
Db 129837 -----AAAAAATTAGACATTTGAACA 129818
Qy 759 nPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProG 779
Db 129817 ACTGGAAGAAAGACATCAAGTGCATTTATCGCTAAATCAATAGTCTTAAAGATCTAA 129758
Qy 779 yIleGlnAlaValAspLeuThrSerThrHisValMetGlySerSerAsnTyrAlaSe 799
Db 129757 AAGTCCAAATGTCCAGAGGTAGGAGGATGATGATGATGATGATGATGATGATGATGATG 129698
Qy 799 rArgGln-ProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsn----- 816
Db 129697 ATCAGTACCTCTCTGAAACAGCCCTTAAGAAACATGTGAAGCTTCAGATCAATACCCAGA 129638
Qy 817 -----GlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAla 833
Db 129637 GGAAAGAGGGCCAAAGTGTCTCAACAAATCCA-----A 129602
Qy 833 erLysLeuCysAspArgArgAsnAla-----GlyGlnValValLeuTyrP 848
Db 129601 TACCTCTGTTAAATCAAGAAATCAATGATGATGATGATGATGATGATGATGATGATGATG 129542
Qy 848 rLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAla 868
Db 129541 TAAAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 129488
Qy 868 erPheProAsnTyrGlyThrSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 129487 ATCTGCAAAATCAGAAATTCATCAGAGAAGAT-CAGGTAGAACAATTCAGCATGCAGAAA 129429
Qy 888 lntYrAlaHisAsn----- 892
Db 129428 AACGAGCCAGAGGCTTTGTACACAGGCTTAACCAATCAGGAGTAATTTGATGTAAG 129369
Qy 893 -----GlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSerThrSerThr 908
Db 129368 AAGTTGAAGTTGCAGAGAACGGGCTCTGAGTGCCATGCTATCAATCCAAATCCATGCTCTG 129309
Qy 908 lY-----LysIleProLeuThrPheGluAspLeuSerArgH 920
Db 129308 GAAGATCAATAAGTCCAAACTTTAAAGTGCCCACTGCTTCACTCAAGAGCAGAAC 129249
Qy 920 isGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlys 940
Db 129248 GCCAACTAGATGACGAGCTCAG-----AGTG 129222
Qy 940 erLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrL 960

Db 129221 CACTTTGGAAAATCAACGATTT-----TCAGCAAGATCCAGA 129183
Qy 960 ysLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluA 980
Db 129182 AGTCTAAATCTTCAGATGTTTTCATGATGCAAAATACCAAGAGAGG 129123
Qy 980 lAlaLeuAsnSerGlyMetPheSerAlaLysTrpAsn-----AlaLeuGlnLeuGlySerV 998
Db 129122 AAAGACTTCAAGTCTTGTCTGCAAGATCAAAATTTCTGCATATATCAATGAAGTCAA 129063
Qy 998 alSerSerAlaAspPhe----- 1004
Db 129062 ATGCTTCTATGACAGATGTCAACACTGATGACAATGAACAGGGGGACAGAGTAGCAAGTG 129003
Qy 1005 -----LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgG 1018
Db 129002 CAGCATCAGTAACATCTGATATTTTACCAAAACAGTATCTCTGCAAAATCCAGGCAT 128943
Qy 1018 lLysGlyLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrA 1038
Db 128942 CCAAGTGTTCGAAGTCCGAACCCGAAGAGCCTGATGATAAAGAAGATTTATGAACATAG 128883
Qy 1038 snLysAsnProAlaAspPheThrThrIleSer 1048
Db 128882 ATAGGTCTCAAGTCAATTTTCAGTAAATATCT 128851
RESULT 14
AC009256/c 184271 bp DNA linear INV 13-DEC-2001
LOCUS Drosophila melanogaster, chromosome X, region 14A-14B, BAC clone
DEFINITION BACR03G18, complete sequence.
ACCESSION AC009256
VERSION AC009256.8 GI:17646883
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 184271)
AUTHORS Ceiliker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Fierres,S., Frise,B., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunco,J.,
Paclet,J., Paragay,V., Park,S., Patel,S., Pfeiffer,B., Pflieger,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome X, region 14A-14B
Unpublished (1998)
TITLE 2 (bases 1 to 184271)
JOURNAL
REFERENCE
AUTHORS Ceiliker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champs,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paclet,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Dec 13, 2001 this sequence version replaced gi:5978361.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdgp@fruitfly.berkeley.edu.

FEATURES

source

```

1. 184271
   /organism="Drosophila melanogaster"
   /mol_type="genomic DNA"
   /strain="y; cn bw sp"
   /db_xref="taxon:7227"
   /chromosome="X"
   /map="14A-14B"
   /clone="BACR03G18 (D892)"
   /clone_lib="RP11-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, Partial 3c0r1 in
pBACE3.6)"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      0.0152      Length:      184271
Score:          233.00      Matches:      241
Percent Similarity: 36.00%      Conservative: 196
Best Local Similarity: 19.85%      Mismatches: 503
Query Match:      4.22%      Indels:      280
DB:              3          Gaps:      48

```

US-09-721-114-2 (1-1057) x AC009256 (1-184271)

```

QY 25 ThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAla 44
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 171046 ACATTTCGATTCTCAATCAATAGCAATGACAAATAGTCCATTACAGATTGAGTGCA 170987
QY 45 GluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyr 64
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170986 GAA-----GATCAACAATCAAGAGAAACATCAGAA-----170954
QY 65 ValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAsp 84
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170953 -----CTTTCAAAATCTTTACCTCAATTAAACACAGAGAAAGCATTTTCATTTCAA 170903
QY 85 GlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArg 104
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170902 GAATCTAGTCGACAAACAATCAATGACAGAGACCG-----170864
QY 105 ArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaPro 124
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170863 ---TGGACACTATCAACATCTCTCTCAACAAGTTCAAGACAAACAATACATTT---170810
QY 125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170809 -----TCGAGTCAGAGTGAATGAAGACAAAACTCTCTAGAGGTACAGG 170762
QY 145 SerThrPheValProAlaSerValGlySerGlnLysValSerProSerThr-----161
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170761 ACATGTCGATTCTCAATCAATT---ACAATGACAAATAGTCCACTTACAGATTTGAGT 170705
QY 162 -----GlnSerSerGlnGlyLysAsnAlaAspArgSer---ThrLeuProLysSer 177
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170704 GCAGAGATCAACAATCAAGAGAAACATCAGACTTTCAAAATCTTTACCTCAATTA 170645
QY 178 ValGlnGlyLysAsn-----AspSerLysCysAsnAlaProSerGlyLysAsn 193
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170644 ACAACAGAAGAAGCATTTTCATTTCAGATCTTAGTGCAGACAAATCAATGACAGA 170585
QY 194 GlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsn 213
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170584 AGAGCCGTGACACATATCAACTCTCTTC-CTCAACAAGTCTCAAGACGAAACATTACAT 170526
QY 214 TyrAspValAlaAlaAsnValSerGluAspAsnThrSerVal---AspValGlyAlaLeu 232
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 170525 TTTCGAGTCAGAGTGTAAATAT---GAAGACAAAACCTTCTTAGAGGATACAAAGGACACTG 170469
QY 233 ProGluValProGlnIleThrTrpHis-----IleGluValAsnGlyAlaAsp 248
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170468 TCGATTCTTCATCAATCAATAGCAATAGTCCACTTACAGATTGAGTGCAGAGAT 170409
QY 249 Gln-----ProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170408 CAACAACATCAAGAGGAAACATCAGAACTTTCAAAATCTTTACCTCAATTAACAACAGAA 170349
QY 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170348 GAAGCATTTTCATTCAGAATCT---AGTGCAGAACAAATCAATCAATGACAGAGCCG 170292
QY 287 AsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThr 304
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170291 TGGACACTATCA---ACATCTCTCTCAACAAGTTCAAGACAAACATTAATTTTCG 170235
QY 305 -----LysAspProLysProValSerGlyGln 313
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170234 AGTCAGAGTGTATATGAAGACAAAACCTTCTTAGAGGATACAGAGCACTGTCGATTCT 170175
QY 314 LysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerLys 333
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170174 CAATCAATTAATGACAAATAGTCCACTTACAGATTGAG-TGCAGAGATCAACAA 170116
QY 334 SerLysArg-----LysThrAspLysLysLeuMetLysLysGlnHis 348
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170115 TCAGAGGAAACATCAGAACTTTCAAAATCTTTACCTCAATTAACAACAGAGAAAGCAT 170056
QY 349 -----SerLysLysArgThrAlaGlnAlaAspVal 358
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 170055 TTCATTTCAGAGATCTAGTCAGACAAACAAATCAATGACAGAGCCGTGGCACTATC 169996
QY 359 SerAspAlaLysLeuCysArgLysProLysValArgLeuLeuSerGluIleLe 378
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169995 AACATCTCTCTCAACAAGTTCAAGACAAAC--ATTACATTTTCGAGTCAGAGTGTA 169938
QY 379 AsnAlaAsnGln-----ValGluAspSerArg-----SerAspGluValHis 392
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169937 AATGAAGACAAACCTTCTTAGAGGATACAGGACACTGTCGATTTCATCAATTACA 169878
QY 393 ArgGluAsnAla-----AlaAspProCysGluAspArgSerThrIleProValPro 410
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169877 ATGACAAATAGTCCACTTACAGATTGAGTCAGAGATCAACAAATCAAGAGGAAACA 169818
QY 411 MetGluValSerMetAspIlePro-----418
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169817 TCAGAACTTTCAAAATCTTTACCTCAATTAACAAGAGAAAGCATTTTCATTCAGAA 169758
QY 419 -----ValSerAsnHisThrValGlyLysAsp-----GlyLeuLys 430
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169757 TCTAGTCAGACAAACATCAATGACAGAGCCGTGGACACTATCAACTCTCTCTCT 169698
QY 431 SerSerLysAsnLysThrLysArgLysTyrSer-----AspValValAspGlySer 448
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169697 CAACAAGTTCAAGACAAACATTTACATTTTCGAGTCAGAGTGTAAATGAAGACAAA 169638
QY 449 SerLeuMetAsnTrpLeuAsnGlyLysLysArgThrGlySerValHisThrVal 468
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169637 TCTCTA-----GAGGATCAAGACACTCTCGATTTCATCAATCA 169596
QY 469 AlaHisProAlaGlyAsnLeuSerAsnLysValThrProThrAlaSerThrGlnHis 488
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169595 -----ACAATGACAAATAGTCCACTTACAGATTGAGTGCAGAGATCA 169551
QY 489 AspAspGluAsnAsp-----ThrGluAsn 496
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169550 ACAATCAAGAGGAAACATCAGAACTTTCAAAATCTTTACCTCAATTAACAAGGAGAA 169491
QY 497 GlyLeu-----AspThrAsnMetHisLysThrAspValCysGlnHisValSerGlu 513
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169490 ACCATTTTCATTCAAGAAATCTAGTCAGAAACAAATCAATCAATCAAGAGCCGTGACA 169431

```


Carlson, J.W., Center, A., Champagne, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferrera, S., Frise, E., Galle, R.F., Gaig, N.S., George, R.A., Gonzalez, M., Hoack, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McInosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleeb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

Sequencing of Drosophila chromosome X, region 14A-14A Unpublished (1998)

2 (bases 1 to 198961)

Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G., Butenhoff, C., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C., Farfan, D.E., Galle, R.F., George, R.A., Harris, N.I., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleeb, J.M., Park, S., Pfeiffer, B., Poon, J., Sequeira, A., Sethi, H., Snir, E., Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rubin, G.M.

Direct Submission

Submitted (27-SEP-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Dec 13, 2001 this sequence version replaced gi:5563419.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

Location/Qualifiers

1..198961

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="X"

/map="14A-14A"

/clone="BACR32K23 (D891)"

/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACe3.6)"

Alignment Scores:

Pred. No.:	0.0167	Length:	198961
Score:	233.00	Matches:	241
Percent Similarity:	36.00%	Conservative:	196
Best Local Similarity:	19.85%	Mismatches:	503
Query Match:	4.22%	Indels:	280
DB:	3	Gaps:	48

US-09-721-114-2 (1-1057) x AC010919 (1-198961)

Qy 25 ThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAla 44

Db 179335 ACATTTCGATTCTCAATCAATAGCAATGACAAATAGTCCTTACAGATTTCAGTGCA 179394

Qy 45 GluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyr 64

Db 179395 GAA-----GATCAACAATCAAGAGAAACATCAGAA-----179427

Qy 65 ValAlaLeuGlnHisLysAspProLysPheCysSerLeuSerArgIlePheHisAsp 84

Db 179428 -----CTTTCAAAATCTTTACCTCAATTAACACAGAAAGCAATTCATTCAA 179478

Qy 85 GlnLysLysCysAspGluHisLysAlaSerSerProPheSerValAlaLysPheArg 104

179479 GAATCTAGTCAGAACAAATCAATATGACAGAGAGCG-----179517

Qy 105 ArgTtpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaPro 124

Db 179518 ---TGACACTATCAACATCTCTCTCTCAACAGAGTTCACAGACAAACAATACATTT---179571

Qy 125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144

Db 179572 -----TCAGTCAGAGTGTAAATGAAGACAAATCTCTCTAGAGATCAAGG 179619

Qy 145 SerThrPheValProAlaSerValGlySerGlnLysValSerProSerThr-----161

Db 179620 ACCTGTGCTGATTTCTCAATCAATTT---ACATGACAAATAGTCCATTCACAGATTGAGT 179676

Qy 162 -----GlnSerSerGlnGlyLysAsnAlaAspArgSer---ThrLeuProLysSer 177

Db 179677 GCAGAGAGATCAACAAATCAAGAGGAAACATCAGAAATCTTCAAAATCTTTACCTCAATTA 179736

Qy 178 ValGlnGluGlyAsn-----AspSerLysCysAsnAlaProSerGlyLysAsn 193

Db 179737 ACAACAGAGAAAGCATTTTCATTTCAAGATCTAGTGCAGAACAAATCAATCAATGACAGA 179796

Qy 194 GlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsn 213

Db 179797 AGAGCGTGGACACTATCAACATCTCTTC-CTCAACAAAGTTCAGAGGACGAACTATCAT 179855

Qy 214 TyrAspValAlaAlaAsnValSerGluAspAsnThrSerVal---AspValGlyAlaLeu 232

Db 179856 TTTCGAGTCAGAGTGTAAATAT---GAACACAAACTCTCTAGAGGATCAGAGCACATG 179912

Qy 233 ProGluValProGlnIleThrTrpHis-----IleGluValAsnGlyAlaAsp 248

Db 179913 TCGATTCTCAATCAATTACATGACAATAGTCCATTCACAGATTTCAGTGCAGAGAT 179972

Qy 249 Gln-----ProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266

Db 179973 CAACAAATCAAGAGGAAACATCAGAACTTTCAAAATCTTTACCTCAATTAACACAGAA 180032

Qy 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286

Db 180033 GAACATTTCAATTCAGAAATCT---AGTCAGACAAACAAATCAATCAATCAGAGAGCG 180089

Qy 287 AnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThr-----304

Db 180090 TGGACACTATCA---ACATCTCTCTCAACAGAGTTCAGAGACAAACATTCATTTTCG 180146

Qy 305 -----LysAspProLysProValSerProValSerGlyGln 313

Db 180147 AGTCAGAGTGTATATGAGACAAACTCTCTAGAGGATCAAGGACACTCTCGATTCT 180206

Qy 314 LysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerLys 333

Db 180207 CAATCAATTACATTCAGCAATAGTCCACTTACAGATTTCAG-TGCAGAGATCAACAAA 180265

Qy 334 SerLysArg-----LysThrAspLysLysLeuMetLysLysGlnGlnHis 348

Db 180266 TCAGAGGAAACATCAGAACTTTCAAAATCTTACTCAATTAACACAGAGAAAGAT 180325

Qy 349 -----SerLysLysArgThrAlaGlnAlaAspVal 358

Db 180326 TTCATTTCAAGAACTCTAGTCAGACAAATCAATGACAGAGAGCGCTGCAGACTATC 180385

Qy 359 SerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIle 378

Db 180386 AACATCTCTCTCAACAGAGTTCAGAGCAAC---ATTACATTTTCGAGTCAGAGTGA 180443

Qy 379 AsnAlaAsnGln-----ValGluAspSerArg-----SerAspGluValHis 392

Db 180444 ATGAGACAAACTCTCTAGAGGATCAGAGCACTGTCTGATTTCTCAATCAATATCA 180503

Qy 393 ArgGluAsnAla-----AlaAspProCysGluAspAspArgSerThrIleProValPro 410

Db 180504 ATGACAAATAGTCCATTACAGATTGAGTCGAGAGATCAAAACAAATCAAGAGAAACA 180563
Qy 411 MetGluValSerMetAspIlePro----- 418
Db 180564 TCAGAACTTTCAAATCTTTACCTCAATTACACAGAGAAAGCAATTTCAATTTCAAGAA 180623
Qy 419 -----ValSerAsnHisThrValGlyGluAsp-----GlyLeuLys 430
Db 180624 TCTAGTCAGAGACAAATCAAAATGACAGAGAGCGGTGGACATATCAACATCTCTCTCT 180683
Qy 431 SerSerLysAsnLysThrLysArgLysTyrSer-----AspValValAspAspGlySer 448
Db 180684 CAACACAGTTCAAGACAAACATACATTTTCGAGTCAGAGTGTAATGACAGACAAACT 180743
Qy 449 SerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisThrVal 468
Db 180744 TCTCTA-----GAGGATCAAGGACACTGTGCGATTTCTCAATCAATA 180785
Qy 469 AlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHis 488
Db 180786 -----ACATGACAAATAGTCACATTACAGATTTGAGTGGCAGAGATCAA 180830
Qy 489 AspAspGluAsnAsp-----ThrGluAsn 496
Db 180831 ACAAAATCAAGAGAAACATCAGAACTTTCAAATCTTTTACCTCAATTAAACAACGGAAGAA 180890
Qy 497 GlyLeu-----AspThrAsnMetHisLysThrAspValCysGlnHisValSerGlu 513
Db 180891 AGCATTTTCATTTCAAGAACTAGTCAGACAAACAAATCAAAATCAGACAGACCGGTGGACA 180950
Qy 514 IleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGlyLysThr 533
Db 180951 CTATCAACATCTCTTTCAACACAGTTCAAGACAAACATATATTTTCGAGTCAGAGT 181010
Qy 534 HisSerAlaAlaSerThrLysTyGlyGlyGluSerThrArgAsn---GlyGlnAsnIle 552
Db 181011 GTAAATGAAGACAAAACTTCTTAGGAAATCAAGGACACTGTGCGATTTCTCAATCAATA 181070
Qy 553 HisVal-----LeuSerAlaGluAspGlnCysGlnMetGluThr 565
Db 181071 ACAATGACAAATAGTCCACTTACAGATTGTAGTCGAGAGATCAACAAATCAAGAG--- 181127
Qy 566 GluAsnSerValLeuSerHisSerAlaLysValSerProAlaGluHisAspIleGlnIle 585
Db 181128 GAAGCATCAGAACTTTCAAATCTTTACCTCAATTAACACAGAGAAAGCAATTTCTATT 181187
Qy 586 MetSerAspLeuHisGluGln-----LysAsnGlnHisGluArg----- 593
Db 181188 CAAGAATCTAGTCAGAGAAACAAATCAATGACAGAGAGCGGTGGACATATCAACATCT 181247
Qy 594 LeuProLysLysLysLysGlnLysLeu-----GluValThrArgGlu 608
Db 181248 CTTCCTCAACACAGTTCAAGACAAACATATATTTTCGAGTCAGAGTGTAATGAGAC 181307
Qy 609 LysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAla----- 625
Db 181308 AAAACTTCTTAGGAGATACAGGACACTGTGCGATTTCTCAATCAATCAATAAACAAT 181367
Qy 626 -----LysAsnGlnHisGluArg----- 631
Db 181368 AGTCTCTTCAGATTTGAGTCGAGAGATCAACAAATCAAGAGAAACATCAGAACTT 181427
Qy 632 -----GlnLeuMetThrGluThrAsp----- 638
Db 181428 TCAAAATCTTTACCTCAATTAACACAGAGAGAAAGCAATTTTCATTTCAAGAACTTAGTGCA 181487
Qy 639 -----Cys 639
Db 181488 GAACAAATCAATGACAGAGAGCGGTGGACATATCAACATCTTCTTCTCAACAAAGT 181547
Qy 640 SerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspCysValIleValala 659
Db 181548 TCAAAGACAAACATATATTTTCGAGTCAGAGTGTAATGAGACAAACAACTTCTCTAGAG 181607

Qy 660 AlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnLysSer 679
Db 181608 GATACAGGACACTGTGCGATTTCTCAATCAATCAATCAATGACAAATAGTCCACTTACAGAT 181667
Qy 680 Leu-----AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAla----- 694
Db 181668 TTGAGTCGAGAAGATCAAAACAAATCAAGAGAAACATCAGAACTTTCAAATCTTTTACCT 181727
Qy 695 ---LeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGln 713
Db 181728 CAATTACACAGAGAAAGCAATTTCAATTCAGAACTTAGTCGAGACAAATCAATCAATG 181787
Qy 714 ThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHis 733
Db 181788 ACA-----GAAGACCGGTGGACACTATCAACATCTTCTTCTCAATCAAGTTC- 181834
Qy 734 HisAspAspGlnTyrIleAlaGluAlaPro-----ThrGluHisTrpGlyArgLys 750
Db 181835 AAAGACAAACATATATTTTCGAGTCAGAGTAAATGAGACAAACTTCTCTAGGG-- 181892
Qy 751 AspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAla 770
Db 181893 GATACAGGACACTGTGCGATTTCTCAATCAAGAGAAATTAACAAATAGT----- 181940
Qy 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 181941 -----CTACTTACAGATTGAGTCGAGAGATCAACAAATCAAGAG--- 181982
Qy 791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAsp----- 808
Db 181983 -----GAAACATCAGAACTTTCAAATCTTTACCTCAATTAACACAGAGAAAGCAAT 182036
Qy 809 ArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAla 828
Db 182037 TCAATTCAGAAATCTGTGCGAGACAAATCAAAATGACAGAGCGGTGGACATATCA 182096
Qy 829 ThrMet-----GluAlaSerLysLeuCysAspArgAsnAlaGlyGlnValValLeu 846
Db 182097 ACATCTCTTCTCAACCAAGTTCAAGACAAACATTACATTTTCGAGTCAGAGTGATAT 182156
Qy 847 TyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeu 866
Db 182157 GAAGACAAACCTTCTCTAGAGGATACAGGACACTGTGCGATTTCTCAATCAATCAATG 182216
Qy 867 AlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsn 886
Db 182217 ACAATAGTCCACTTACAGATTTGAGTCGAGAGATCAACA----- 182258
Qy 887 SerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeu 906
Db 182259 -----AATCAAGAGAAACATCAGAACTTTCAAATCTTTTACCTCAATTA 182303
Qy 907 AsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHis 926
Db 182304 ACAACAGAGAAAGCAATTTCAATTCAGAAATCTAGTCGAGACAAATCAATCAATGACGAA 182363
Qy 927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
Db 182364 AAGCGGTGGACACCATCAACATCTCTTCTCAACCAAGTTCAAGACAAACATTAATTT 182423
Qy 947 AlaAsnTrpSer-----GluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSer 964
Db 182424 TCGAGTCAGAGTGATATGAAGACAA-AACTTCTCTAGAGGATACAGGACACTGTGCGAT 182482
Qy 965 ThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGly 984
Db 182483 TTCTCATCAAGAGAAATTAACAAATAGTCTTACATTCAGATTTGAGTCGAGAGATCAAC 182542
Qy 985 MetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPhe 1004
Db 182543 AAA-----TCAAGAGAAACATCAGAACTTTCAAAATCTTTT 182578

```
Qy 1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024
Db 182579 ACCTCAATTAAACAACAGC----- 182596
Qy 1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn-----Pro 1041
Db 182597 ---AGAAAGCATTTTCATTTTCAGGAATCTAGTGCAGAACAAATCAATGACAGAGAGCC 182653
Qy 1042 AlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAsp 1055
Db 182654 GTGGACACTATCAACATCTCTTCTCAACAACAAGTTCAAGAC 182695
```

Search completed: May 17, 2004, 08:43:32
Job time : 11717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2004, 01:43:21 ; Search time 955 Seconds
(without alignments)
4701.932 Million cell updates/sec

Title: US-09-721-114-2

Perfect score: 5526

Sequence: 1 MEIVAVDQEGARVVGTCNML.....NKRPADFTTISNDNRYMDYR 1057

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dgapop 6.0 , Delxext 7.0

Searched: 337363 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/Cgn2_1/USPFC0001/US09721114/runat 14052004 092250 4074/app query.fasta_1.1223
-DB=N_Geneseq_29Jan04 -QEXT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFCIL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-NO_XMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=126 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04: *
1: geneseq1980s: *
2: geneseq1990s: *
3: geneseq2000s: *
4: geneseq2001as: *
5: geneseq2001bs: *
6: geneseq2002as: *
7: geneseq2003as: *
8: geneseq2003bs: *
9: geneseq2003cs: *
10: geneseq2004as: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5526	100.0	3896	7 ABQ77908	Abq77908 Rice OsEMF1
2	5526	100.0	4310	6 ABA93084	Abq93084 Oryza sat
3	5191	93.9	9455	6 ABA93085	Abq93085 Oryza sat
4	393	7.1	3291	7 ABZ42128	Abz42128 Arabidops
5	298	5.4	8646	3 AAA14873	Aaa14873 Nucleotid
6	298	5.4	8648	6 AAL43177	Aal43177 Arabidops
7	298	5.4	8648	7 ABQ77909	Abq77909 Arabidops
8	298	5.4	17341	3 AAA14872	Aaa14872 Genomic D

9	298	5.4	17341	6 AAL43176	Aal43176 Arabidops
10	223.5	4.0	5286	5 AAS73156	Aas73156 DNA encod
11	223	4.0	10452	6 ABL68623	Ab168623 Kidney ca
12	223	4.0	10452	6 ABL68623	Ab168623 Kidney ca
13	223	4.0	11167	9 ADES3841	Ades3841 Human pro
14	222.5	4.0	279	6 ABL73385	Ab173385 Corn tass
15	221	4.0	5954	5 AAS80591	Aas80591 DNA encod
16	221	4.0	6143	5 AAS83843	Aas83843 DNA encod
17	220	4.0	3953	4 ABA66556	Ab665656 Human foe
18	220	4.0	3953	4 ABA66556	Ab665656 Human foe
19	220	4.0	3953	4 ABA66556	Ab665656 Human foe
20	220	4.0	3953	4 ABA66556	Ab665656 Human foe
21	220	4.0	3953	4 ABA66556	Ab665656 Human foe
22	220	4.0	3953	4 ABA66556	Ab665656 Human foe
23	220	4.0	3953	4 ABA66556	Ab665656 Human foe
24	220	4.0	3953	4 ABA66556	Ab665656 Human foe
25	219	4.0	6201	7 ABX71059	Abx71059 Novel hum
26	217	3.9	3998	4 AAH29827	Aah29827 S cerevis
27	214	3.9	5072	2 AAT71315	Aat71315 CDNA enco
28	214	3.9	5073	2 AAT71314	Aat71314 CDNA enco
29	214	3.9	9642	2 AAT71324	Aat71324 Plasmid p
30	214	3.9	10443	2 AAT71324	Aat71324 Plasmid p
31	214	3.9	13414	2 AAT71321	Aat71321 Plasmid p
32	212.5	3.8	6816	7 ACA19870	Aca19870 Prokaryot
33	211.5	3.8	6783	7 ABT14863	Abt14863 Pathogen
34	211	3.8	7474	2 AAT71320	Aat71320 Plasmid p
35	209	3.8	6813	7 ACF74294	Acf74294 Staphyloc
36	209	3.8	6852	7 ABZ22900	Abz22900 Staphyloc
37	209	3.8	9807	4 ABL06949	Ab106949 Drosophil
38	207	3.7	5495	7 ACA46979	Aca46979 Prokaryot
39	207	3.7	5580	7 ABT14982	Abt14982 Pathogen
40	206.5	3.7	9606	4 ABL12165	Ab112165 Drosophil
41	206.5	3.7	12039	4 ABL12164	Ab112164 Drosophil
42	204.5	3.7	10219	4 ABL20122	Ab120122 Drosophil
43	204	3.7	5355	4 ABL29287	Ab129287 Drosophil
44	204	3.7	8307	4 ABL23161	Ab123161 Drosophil
45	204	3.7	8431	9 ADD29688	Add29688 Human tum

ALIGNMENTS

RESULT 1
ABQ77908
ID ABQ77908 standard; cDNA; 3896 BP.

XX AC ABQ77908;

XX DT 24-JAN-2003 (first entry)

XX DE Rice OsEMF1-encoding cDNA, SEQ ID NO:1.

XX KW Rice; OsEMF1; EMF; embryonic flower; plant; floral repressor;
XX KW reproductive development; flower development; transgenic plant;
XX KW antisense suppression; transgenic; reciprocal negative interaction;
XX KW flower meristem identity gene; flowering time; shoot development;
XX KW seed yield; agriculture; gene; ss.

XX OS Oryza sativa.

XX FH Key Location/Qualifiers

XX CDS 241..3414

XX FT /*tag= a

XX FT /product= "OsEMF1 protein"

XX FN WO200280659-A1.

XX PD 17-OCT-2002.

XX PF 05-APR-2002; 2002WO-US012675.

XX PR 06-APR-2001; 2001US-00828068.

XX PA (REGC) UNIV CALIFORNIA.

XX Moon Y, Chen L, Sung ZR;
 XX WPI; 2003-046831/04.
 DR P-PSDB; ABB99878.
 XX
 PT New isolated OsEMF1 nucleic acid molecule and encoded polypeptide, useful
 PT for plant genetic engineering, in particular controlling reproductive
 PT development in rice.
 XX
 PS Claim 3; Fig 1; 47pp; English.

XX The invention relates to a rice EMF (embryonic flower) gene designated
 CC OsEMF1 (CDNA given in AB077908) and its encoded protein (ABB99878). The
 CC OsEMF1 protein has a molecular weight of 116.4 kD and exhibits 37%
 CC homology and 20% identity with Arabidopsis thaliana EMF1 protein
 CC (ABB99879). OsEMF1, like other EMF gene products, acts as a floral
 CC repressor, suppressing the transition from vegetative growth to
 CC reproductive development. It also delays the inflorescence to flower
 CC transition, indicating that there is a reciprocal negative interaction
 CC between OsEMF and flower meristem identity genes. The invention also
 CC encompasses a transgenic plant comprising an expression cassette
 CC containing an OsEMF1 nucleic acid (particularly in an antisense
 CC orientation) under the control of a plant promoter. OsEMF1 nucleic acids
 CC and proteins may be used for controlling reproductive development in
 CC plants, particularly monocotyledonous plants and especially rice. In
 CC particular, OsEMF1 nucleic acids and proteins may be used to control
 CC flowering time, shoot development and seed yield. Controlling or
 CC inhibiting the expression of genes which mediate these processes enables
 CC new varieties of rice with different flowering times and seed yield to be
 CC developed. The present sequence represents OsEMF1 cDNA. Note: The present
 CC sequence is described as SEQ ID NO.1 in the claims and sequence listing,
 CC but the sequence referred to as SEQ ID NO.1 in the examples (not shown in
 CC the specification) is described as an Arabidopsis thaliana EMF1 genomic
 CC clone
 XX
 SQ Sequence 3896 BP; 1176 A; 791 C; 935 G; 994 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3896
 Score: 5526.00 Matches: 1057
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-09-721-114-2 (1-1057) x ABQ77908 (1-3896)
 Qy 1 MetGluLeuValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
 Db 241 ATGGAGATTGTTCAGTAGATCAGGAGGAGCTCGTCTTGTGGACAACTGTATGCTT 300
 Qy 21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
 Db 301 GCTCGTGGTGAAGTGTGTGTAGCCGACGCTGTGTGGAGCTGACGAGCGCTCGTCAG 360
 Qy 41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
 Db 361 GATGACCGCTGAGCTGGTGTAGACGACCGGCACACACCAATGCGAGCATTTCTCC 420
 Qy 61 IleArgGlyTyrValAlaLeuLeuGlnHisLysAspProLysPheCysSerLeuSerArg 80
 Db 421 ATAAAGAGGGTATGTGTCTCTCTCAGAGAAGATCCAAATTCCTCTCTCTATCTCGG 480
 Qy 81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
 Db 481 ATTTTCCATGCCAGAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCAATTTCTGTA 540
 Qy 101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
 Db 541 GCAAAAGTTTCACGATGGGATTGCTCGAAGTGTCTGGATAAGTTGAAATTCAGATAAT 600
 Qy 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140

601 GGACAGCAGCACCAGAACTCTTCCGCAAGACAGAAAGGACAGTGGTGTCTCCATC 660
 Qy 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
 Db 561 ACATTGTTCGAGGACATTTTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 720
 Qy 161 ThrGlnSerSerGlnGlyAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
 Db 721 ACACAATCATCTCAAGGAGGAAGAAATCTGATAGATCACTCTTCCAAAGAGTGTGCAAGAA 780
 Qy 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyValAlaGluAlaAsnThr 200
 Db 781 GGCAATGACTCCAAATGCAATCGGCTTCTGCAAGAAATGGAGCTGCTGAGGCCAATACT 840
 Qy 201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnLysAspValAlaAlaAsnVal 220
 Db 841 GATTCCACCAATCAAGATTTTGCAGGGCCAGCCCAAAATATGATGTGGCAGCAAAATGTC 900
 Qy 221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
 Db 901 TCTGAGGACACACTTCTGT 960
 Qy 241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
 Db 961 CACATAGAGTAATGTGTGAGATCACTCCATCCACTCCAAAACTTTCTGAAGTGTCT 1020
 Qy 261 LeuLysArgAsnGluAspGlnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
 Db 1021 CTCAAAAGAAATGCAAGATGAAATGCAAAATGCAAGAGACTCTTGTGTCTGAGCAGTGC 1080
 Qy 281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
 Db 1081 AATTTGACCAAGATCTTAACCCATGTCTGGAAGGAACGATGATCAGGTTCTGAGCAG 1140
 Qy 301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
 Db 1141 TGCATTTTGACCAAGATCCGAAACAGTGTCTGGGCAAGAAATGTGACAGATCTGCAAT 1200
 Qy 321 GluProCysGluGluValValLeuLysArgSerSerLysValSerLysArgLysThrAspLys 340
 Db 1201 GAGCATGTGTGAAGAGTTGTCTCAAGAGACTCCAAATCTAAGGGAGAGCGGTAAG 1260
 Qy 341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
 Db 1261 AAGTTGATGAGAAGCAGCAGCAGCAAGAAACGACTGCCCGAGCTGATGTTTCAGAT 1320
 Qy 361 AlaLysLeuCysArgArgLysProLysValAlaArgLeuSerGluIleIleAsnAla 380
 Db 1321 GCARAAGCTTTGTGGAGAAAGCCAAAGAGTGGCGCTCTCTATCAGAAATTTAATAGCT 1380
 Qy 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
 Db 1381 ACCAGTTGAGGATTTAGAAAGTCAAGAGTTCATCGTGAATAATGCGCTGATCCCTGT 1440
 Qy 401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
 Db 1441 GAGGATGATAGAAAGTACCATCCCGTCCGATGGAAAGTAAGCATGGATATCTCTGTAGC 1500
 Qy 421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
 Db 1501 ACCATACATAGTGGAGAGATGGTTAAATCAAGTAAGAACACAGCAAAACCCAAATAC 1560
 Qy 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysValArg 460
 Db 1561 TCTGATGTTGTAGATGATGATCATCTATTAATGAATGGCTGGCTGAATGGAAGAAAGAAAGA 1620
 Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysVal 480
 Db 1621 ACTGGAGTGTGCATCACACACTTGTCTCATCCAGCTGGGAATTTGAGCAACAAAAAGTGTG 1680
 Qy 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLysLeuAspThr 500

Db 1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAAATGCTGTTGACACA 1740
Qy 501 AsnMetHisThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db 1741 AATATGCATTAAGACAGATGCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Qy 521 SerIysGlyIysThrAlaGlyLeuSerIysGlyIysThrHisSerAlaAlaSerThrIys 540
Db 1801 TCAAGGGGAAAAACAGCGGGTTTCAGTAAGGGGAAAAACATTCAGCTGCTAGTACCAAA 1860
Qy 541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db 1861 TATGGTGGTGAAGACACAGAAATGTCAGAAACATACATGCTACTGCGGCGAGAGATCAA 1920
Qy 561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaIysValSerProAlaGlu 580
Db 1921 TCCAGATGGAAACCGAAATCTGTTCTGAGTCACTCGGCAAGGTTTCTCCAGCTGAG 1980
Qy 581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLys 600
Db 1981 CATGATATCCAAATATGTCCTGACCTTCATGACGAGAGTCTACCCAAAGAAAAGAG 2040
Qy 601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db 2041 CAAAAAATGGAAGTGAAGTCTGCGTGAAGAAACAGACCATGATGATGATGATGATGAT 2100
Qy 621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db 2101 GTGACCTGCTAGCTAAACCAACATGACGAGGAGCTTATGACTGAGACTGATGTCT 2160
Qy 641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspCysValIleValAlaAla 660
Db 2161 GACATCAACCGTATTCAATCCAGACAACTGCTGATGATGATGATGATGATGATGAT 2220
Qy 661 LysAspGlySerAspThrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db 2221 AAGGATGGTTCAGATTAATGATCAAGTGTGTTGACACATTAATCCCAACAGAGTCTCT 2280
Qy 681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db 2281 GCATCCCAAGTACACAGAGAGGTACAGGCTCATTTGGCATTTGACACACACAGAGTCT 2340
Qy 701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrHisLeuArgMetGluGlu 720
Db 2341 CCACATCCTCAGAACTTTTCACTGCTACTCAGGAAACAGCAGACACATTTGCGGAGTGA 2400
Qy 721 MetValThrIleAlaAlaSerSerProLeuPheSerHisAspAspGlnTyrIleAla 740
Db 2401 ATGCTCACTATTGCTGACGCTCACCCTATTTTTCATCATGATGATGATGATGATGAT 2460
Qy 741 GluAlaProThrGluHisTyrGlyArgLysAspAlaLysLysLeuThrTyrGluGlnPhe 760
Db 2461 GAAGCACCACCTGAACATTTGGGCGCGTAAGCAGCAGCAAGAGCTAAAGTGGAGCAAT 2520
Qy 761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db 2521 AAGGCCCACTACAGAAATTTCCAGCAGCAGCAATGCTGCTCAATTTAGACCTGGTATC 2580
Qy 781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db 2581 CAACAGTGTGACTTGACTTCTACTCATGTCATGGAATCTTCCAGCAATTAATGATCTCGC 2640
Qy 801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db 2641 CAACCACTTAATGGCCACTGGACCGCTATGCTGAAAGAGCGGTAAACAGGTCCTATGCA 2700
Qy 821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
Db 2701 AGAAATTTCCAGACCAATAGCAACCAATGAGCGAGTAAAGTTATGTCATCGAGAAAT 2760
Qy 841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
Db 2761 GCTGGACAAGTAGTCTTGATCTCTAAAGAAATCCATGCTCGCAGCATCTTCTCAGAAATG 2820

Qy 861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
Db 2821 ATGGATCCATCAACATTAAGCAAGCTTCCCACACTATGGAACCTTAGCAGAAACCAAGATG 2880
Qy 881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
Db 2881 GAGTCTCAACTTCATATATCTCAGTATGACATATACATACATACAGATCAACAGACCA 2940
Qy 901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluLeuAspLeuSerArgHis 920
Db 2941 TCATATGGCAGTAACCTGAATGGAAGATTCCTATTCATTCGAGACTTATCACGSCAT 3000
Qy 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
Db 3001 CAGTCGATGATCTGCACAGACCTTTACGCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Qy 941 LeuLeuGlnLysGluIleAlaAsnTyrSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
Db 3061 TTCTCTCAGAGGAATTCGAAACTGGTCGAGACACTGTGGCACACAACTCTGGTTATAAG 3120
Qy 961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
Db 3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAAGTGAACAGAAAGGACATTTTGAAGCC 3180
Qy 981 LeuAsnSerGlyMetPheSerAlaLysTyrAsnAlaLeuGlnLeuGlySerValSerSer 1000
Db 3181 CTGAATTCCTGGAATGTTTTCAGCAAAATGGAATGCAATTCAGTTCGGTTCGTAGCTCC 3240
Qy 1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerThrThrArgGlyLysGly 1020
Db 3241 AGTGCAGATTTTATCAGCGAGAAACAGCATAGCTCAATCTTGGACCAAGCAAGGGT 3300
Qy 1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
Db 3301 AAAATGTTTCATCCCTGGATCGTTTGTGAGACAGATATCTGTATTACTACAGAAC 3360
Qy 1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
Db 3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGATTAACCGC 3411

RESULT 2
ABA93084
ID ABA93084 standard; cDNA; 4310 BP.
XX
AC ABA93084;
XX
DT 11-APR-2002 (first entry)
XX
DS Oryza sativa brassinosteroid response related protein encoding cDNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
XX signal transfer system; brassinosteroid hormone; growth promotion;
XX increased yield; quality improvement; ripeness promotion;
XX stress relaxation; chemical resistance; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 655..3828
FT /tag= a
FT /product= "brassinosteroid response related protein"
XX
FN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PP 19-MAY-2000; 2000JP-00149106.
XX
PP 19-MAY-2000; 2000JP-00149106.
XX
PP (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKEI TOKUTEI SANGYO GIJUTSU.
PA

XX WPI; 2002-135739/18.
DR P-PSDB; ABB05428.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
XX
PS Disclosure; Page 6-10; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence encodes a
CC brassinosteroid response related protein isolated from rice (*Oryza*
CC *sativa*), from the present invention
XX
SQ Sequence 4310 BP; 1237 A; 986 C; 1022 G; 1065 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 4310
Score: 5526.00 Matches: 1057
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-721-114-2 (1-1057) x ABA93084 (1-4310)
QY 1 MetGluLeuValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
DB 655 ATGGAGATTGTTCAGTAGATCAGGAGGAGCTGCTGTTTGGGAGCACTGTATGCTT 714
QY 21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
DB 715 GCTCGTGGTGAAGTCTGTGTAGCCGAGTGTGGAGCTGACAGGACGCTCGTCAG 774
QY 41 AspAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
DB 775 GATCAGCGCTGAAGCTGTGTAGACCAACCGGACCAACCAACCACTCCGAGCAATTCCTCC 834
QY 61 IleArgGlyTyrValAlaLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
DB 835 ATAGAGGGTATGTGTCTCTTCAGAGAGAGTCCAAAATCTCTCTCTATCTCGG 894
QY 81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerProPheSerVal 100
DB 895 ATTTTCCATGACCAAAAATGTGATCAACACAAAGCTAGTTCAGGCCCATTTCTGTA 954
QY 101 AlalysPheArgArgTropAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
DB 955 GCAAAGTTTCACCATGGATGCTCGAAGTGTCTGGATAGTTCGAAAGCTTCAGATAT 1014
QY 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerile 140
DB 1015 GGACAGCACCAAGAACTCTTCCGAAAGCAGAAATGACAAAGTGTGCTCCATC 1074
QY 141 ThrPheValArgSerThrPheValProHisSerValGlySerGlnLysValSerProSer 160
DB 1075 ACATTGTTCGGAGACATTTTGTGCTGTAGTGTGTTCCCAAAAGTGTCTCCTAGC 1134
QY 161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
DB 1135 ACACANTCATCTCAAGGGAGAGATGCTGTATAGATCACTCTTCCAAAGAGTGTGCAAGAA 1194
QY 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
DB 1195 GGCATGACTCCAAATGCAATGCGCTCTTGGCAGAGATGGAGTGTCTGGAGGCCAAATCT 1254
QY 201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
DB 1255 GATTCCCAATGAAGATTTGCAAGGCGGCGCCCAAAATTTATGATGTGGCAGCAATGTC 1314

QY 221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
DB 1315 TCTGAGGACACACATCTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCGACATTACATGG 1374
QY 241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
DB 1375 CACATAGAAGTAATATGTTGTCAGATCAACTCCATCCACTCCAAAACCTTTCTGAAGTGGTC 1434
QY 261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
DB 1435 CTCAAAAGAAATGAGATGAAAATGGAAAACCTGAAGAGACTTTGTTGCTGAGCAGTGC 1494
QY 281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
DB 1495 AATTGACCAAGATCTCTAAACCCAAATGCTGGAAGGAACGATGATCAGGTTTCTGAGCAG 1554
QY 301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
DB 1555 TCAATTTTGACCAAGATCCGAAACCGATGCTGGGCGAAGATGTGAGCAGATCTGCAAT 1614
QY 321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
DB 1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAGCTCCAAATCTAAGAGGAAGACGATAG 1674
QY 341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
DB 1675 AAGTTGATGAAGACGAGCAGCAGCAAGCACTGCCCGAGCTGATGTTTCAGAT 1734
QY 361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleLeuAsnAla 380
DB 1735 GCARAGCTTTGCGGAGAAAGCCAAAAGAGTGGCGCTTCTATCAGAAATTAATAATGCT 1794
QY 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
DB 1795 AACCAAGTTGAGGATCTAGAAAGTTCATCGTGAAGATGCGCCCTGATCCCTGT 1854
QY 401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
DB 1855 GAGATGATAGAGTACCATCCCGTCCGATGGAAGTAAGCATGATATTCCTGTATGC 1914
QY 421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
DB 1915 AACCATCAGTGGGAGAGATGGGTAAATCAAGTAAGAACACAGACAAAACGCAAAATAC 1974
QY 441 SerAspValValAspAspGlySerSerLeuMetAsnTyrLeuAsnGlyLysLysLysArg 460
DB 1975 TCTGATGTTGTAGATGATGGATCATCATTAATGACTGCTGATGGAAGAAAGAAAGA 2034
QY 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
DB 2035 ACTGGAAGTGTGCATCACACAGTGTCTCATCCAGCTGGGAATTTGAGCAACAAAAGTG 2094
QY 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
DB 2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGATGATGATGATACTGAAAATGTTTGCACA 2154
QY 501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
DB 2155 AATATCATTAAGACAGATGCTGTGTCAGCATGTATCAGAAATCTCCACACAGAGTGTCA 2214
QY 521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
DB 2215 TCAAGGGGAAAACACGCGGTTTGATGAAGGGGAAAACATTCAGCTGCTAGTACAAA 2274
QY 541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
DB 2275 TATGTTGTTGAAGACACAGAAATGTGTCAGACATACATGATCTCAGCGCAGAAAGTCAA 2334
QY 561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
DB 2335 TGCCAGATGGAACCCGAAAACCTCTGTCTGAGTCACTCGGCAAGAGTTTCTCCAGCTGAG 2394

Db	3475	TTCTCTCAGAGGAAATTTGCAAACTGGTCCGAGAACTGTGGCACACATCTGGTTATAAG	3534
Qy	961	LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla	980
Db	3535	TTAGGAGTGTCAACAGGAATAACATCCGATCAGATGAACAGAAAGGAACATTTTGAAGCC	3594
Qy	981	LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer	1000
Db	3595	CTGAATTCGGAAATGTTTTCAGCAAAATGAATGCATTGCAGTTGGGTTCTGTAGCTCC	3654
Qy	1001	SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly	1020
Db	3655	AGTGCAGATTTTTCATCAGCGAGGACAGCATAGCTCAATCTTGGACCAGAGCAAGGGT	3714
Qy	1021	LysMetValHisProLeuAspArgPheValArgGlnAsyIleCysIleThrAsnLysAsn	1040
Db	3715	AAAATGGTTTCATCCCTTGATCGGTTTGTGAGACAGGATATCTGTATAACTATCAAGAAC	3774
Qy	1041	ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg	1057
Db	3775	CCAGCTGATTTTACTACATCAGTAACGATACGAGTATATCGATTACCGC	3825
RESULT 3			
ABA93085			
ID ID ABA93085 standard; DNA; 9455 BP.			
XX	AC	ABA93085;	
XX	DT	11-APR-2002 (first entry)	
XX	XX	Oryza sativa brassinosteroid response related protein related DNA.	
XX	XX	Oryza sativa; rice; plant; brassinosteroid response; control;	
KW	KW	signal transfer system; brassinosteroid hormone; growth promotion;	
KW	KW	increased yield; quality improvement; ripeness promotion;	
KW	KW	stress relaxation; chemical resistance; gene, ds.	
XX	OS	Oryza sativa.	
XX	XX	JP2001327287-A.	
PN	PN	27-NOV-2001.	
PD	PD	19-MAY-2000; 2000JP-00149106.	
PF	PF	19-MAY-2000; 2000JP-00149106.	
PR	PR	(NORO) NORINSUISANSHO MOGYO SEIBUTSU SHIGEN.	
XX	XX	(SEIB-) SEIBUTSUKEI TOKUWEI SANGYO GIJUTSU.	
PA	PA	WPI; 2002-135739/18.	
XX	XX	New gene involved in brassinosteroid responses useful for controlling the	
PT	PT	effects such as growth promotion, increased yield, quality improvement,	
PT	PT	ripeness promotion, stress relaxation and chemical resistance.	
XX	XX	Disclosure; Page 13-16; 19pp; Japanese.	
PS	PS	The present invention describes a polynucleotide encoding a plant gene	
XX	XX	which can control the signal transfer system of brassinosteroid hormone.	
CC	CC	The polynucleotide can be used for controlling the effects such as growth	
CC	CC	promotion, increased yield, quality improvement, ripeness promotion,	
CC	CC	stress relaxation and chemical resistance. The present sequence	
CC	CC	represents a DNA sequence related to a brassinosteroid response related	
CC	CC	protein isolated from rice (Oryza sativa), from the present invention	
XX	XX	SQ Sequence 9455 BP; 2571 A; 2001 C; 2090 G; 2793 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.: 0 Length: 9455			
Score: 5191.00 Matches: 1055			
Percent Similarity: 58.06% Conservative: 0			

```

Alignment Scores:
pred. No.:      0
Score:          5191.00
Length:         9455
Matches:        1055
Percent Similarity: 68.06%
Conservative:    0

```

Best Local Similarity: 68.06%				Mismatches: 2			
Query Match: 93.94%				Indels: 495			
DB: 6				Gaps: 3			
US-09-721-114-2 (1-1057) x ABA93085 (1-9455)							
QY	1	MetGluileValAlaValaspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu	20	QY	61	-----	61
Db	4223	ATGGAGATTGTCAGTAGATCAGGAGGAGCTCGTGTGTTGGGACGAACGTATGCTT	4282	Db	5242	TTGTAGTTTATCCTGTCATGCTATTGTTATTATCTTTGAAATTCAAACTGCAATACTTAGA	5301
QY	21	AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln	40	QY	61	-----	61
Db	4283	GCTCGTGTGGAACTGCTGCTGTAGCCGAGTGTGGAGCTGACAGCGACGCTCGTCAG	4342	Db	5302	TTATCTTGAAGGTCTCTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACTCTCC	5361
QY	41	AspAlaAlaGluAlaGlyValaspGluProAlaGlnHisGlnCysGluHisPheSer	60	QY	61	-----	61
Db	4343	GATCAGCGCGCTGAGCTGGTGTAGACGAACCGGCACACCAATGCGAGCATTTCTCC	4402	Db	5362	AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC	5421
QY	61	Ile-----	61	QY	61	-----	61
Db	4403	AT-AAGTAATCAITTTTCTGTATTTTCCAATTCAGTATCGGTTGTGGATGAATAGAA	4461	Db	5422	ACGCAATGTCTATATATATATGTGCCATGCACTTTTATAATGATAATCTTATTTTCTGT	5481
QY	61	-----	61	QY	62	---ArgGlyValAlaLeuLeuGlnLysValLysAspProLysPheCysSerLeuSerArg	80
Db	4462	TCGGCAATGTCATGCCATATTCACATGTTTGTATGGAAGATGATGATGCTGTTTGG	4521	Db	5482	ACRAGAGGTATGTCTCTTCTTCAAGAAGAGATCCAAATCTCTGCTCTATCTCGG	5541
QY	61	-----	61	QY	81	IlePheHisAspGlnLysCysAspGluHisLysAlaSerSerSerProPheSerVal	100
Db	4522	CACAGTTTCTGTTGGCACTTATATGTCATCTGTTTGTAGCATCGTATACACTGGTC	4581	Db	5542	ATTTTCCATGACCAAGAAAAATGTATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA	5601
QY	61	-----	61	QY	101	AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn	120
Db	4582	GACATGCTTATGACTTGTGTTGATGATTTAGGAATCAATACATCCACTACTAGCTCTAT	4641	Db	5602	GCAAAAGTTTCGACGATGGGATGCTCGAAGTCTTGGATAGTTGAAAACCTTCAGATAAT	5661
QY	61	-----	61	QY	121	GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle	140
Db	4642	CTAGCCATGTGAACCTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAATA	4701	Db	5662	GGACAGCACCAAGAACTCTTCCCGCAAGCAGAAATGGCACAAGTATGGTTGCTCCATC	5721
QY	61	-----	61	QY	141	ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer	160
Db	4702	TATAATATTTGCATATATGTTGGTTCATGTAATCTTTTACTCTTACTCTAGTACATTA	4761	Db	5722	ACATTTGTTTCGAGCACTTTTGTGCTGCTAGTTGGTTGGTTCCCAAAAGTGTCTCTAGC	5781
QY	61	-----	61	QY	161	ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu	180
Db	4762	ATATCTCAATGATGAATCTGACACATGATTTGAGTGAGTGCTACACATATGCCATGCT	4821	Db	5782	ACACATATCTCTCAAGGGAAGATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA	5841
QY	61	-----	61	QY	181	GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyLysAlaAlaGluAlaAsnThr	200
Db	4822	ATGTGTGTTCAATAGGTGTTGATCATATTTGTTGTTGTTGGGTGCGCARGCAATTTAT	4881	Db	5842	GGCAATGACTCCAAATGCAATGCGCTTCTGGCAAGATGGAGCTGTGAGGCCAAATACT	5901
QY	61	-----	61	QY	201	AspSerProMetLys-----	205
Db	4882	CAGGCCATGCTGTAGGCTGTAGATATTTGTGTTGTATATTTCTGTGTAACAA	4941	Db	5902	GATTCACCAATGAA-AGGTATGTTAGTAGATGTAGAGCTTTCAAAATTCCTAAGTAGGATTTT	5960
QY	61	-----	61	QY	206	-----	211
Db	4942	GCTGATTACTAATGAATTAACCTTTTGGGGTACACTCATATATTGGGCCCTACATTT	5001	Db	5961	ATTTAAGGTATAGATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC	6020
QY	61	-----	61	QY	212	GlnAsnTyAspValAlaAlaAlaSerGluAspAsnThrSerValAspValGlyAla	231
Db	5002	TGTAATCAATTTTCTTGTGCTGAGGTTACAGATAAAACTTTTATATCATACATGTT	5061	Db	6021	CAAAATTTATGATGTGGCAGCAATGCTCTGAGGACCAACTCTCTGTTGATTTGSGGCT	6080
QY	61	-----	61	QY	232	LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro	251
Db	5062	TACATCCTAGGAGATCTTAGAATGATGTTTCTTCATATTTGCAATTTGCAATTTG	5121	Db	6081	TTACCTGAAGTTCCCCAGATTACATGGCACATAGTAAGTAATGCTGCAGATCAACCTCCA	6140
QY	61	-----	61	QY	252	SerThrProLysLeuSerGluValValLysValAsnGluAspGluAsnGlyLysThr	271
Db	5122	ATAGTCCATTATTTTAAAGCTTTTCAATTTGTTTAGAGATCTAGAGATGATATATAT	5181	Db	6141	TCCACTCCAAAACCTTTCTGAAGTGGTCTCTCAAAAGAAATGAAGATGAAATGAAAACCT	6200
QY	61	-----	61	QY	272	GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly	291
Db	5182	CAACCATAGACTTGTACAGTTTGGTTTAACTTTCTAGAACTTAATTAGATTATTTT	5241	Db	6201	GAGAGACTCTTGTGCTGAGCAGTGCATTTTGACCAAGATCTCTTAACCAATGTCTGGA	6260
QY	61	-----	61	QY	292	LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer	311
Db	5241	-----	61	QY	6261	AAGGAACGTGATCAGGTTGCTGAGCAGTGCATTTTGACCAAGATCCGAACCCAGTGTCT	6320

QY 312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluValValLeuLysArgSer 331
DB 6321 GGGCGAAATGTGAGCAGATCTGCAATGACCGATGTGAAGATGTGTTCTCAAAAGAAC 6380
QY 332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
DB 6381 TCCAAATCTTAAGGAGAGACGGATAGAAAGTTGATGAAGAGCGACGACCAAGAA 6440
QY 352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
DB 6441 CGCACTGCCCGGTGATGTTTCAGATGCAAGCTTTGTGCGGAAAGCCAAAAGGTG 6500
QY 372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
DB 6501 CGGCTTCATCAGAAATTAATAATCTTAACAGGTTGAGATTCAGAGTTCAGAGTT 6560
QY 392 HisArgGlnAsnAlaAlaAspProCysGluAspArgSerThrIleProValProMet 411
DB 6561 CATCGTGAAATGCCGCTGATCCCTGTGAGCATGATAGAAAGTACCACCCCGTCCCGATG 6620
QY 412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
DB 6621 GAAGTAAGCATGGATATCTCTGTTAGCAACCATACAGTGGGAGAGATGGTTAAATCA 6680
QY 432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
DB 6681 AGTAAGAACCAAGACAAACCAATACTCTGATGTTGTAGATGATGATCATCACTTATG 6740
QY 452 AsnThrLeuAsnGlyLysLysArgThrGlySerValHisThrValAlaHisPro 471
DB 6741 AACTGGCTGAATGGAAGAAAGAAAGAAAGTGGAGTGTGATCACACAGTTGTCTATCCA 6800
QY 472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
DB 6801 GCTGGGATTTGACCAACAAAGATGACACCCACTCGGAGTACTCAGCATGATGATGAG 6860
QY 492 AsnAspThrGluAsnGlyLeuAspThrAsnMethHisLysThrAspValCysGlnHisVal 511
DB 6861 AATGATACTGAAATGGTCTTTGACACAAATATGATGAAGACAGATGCTCTCAGCATGTA 6920
QY 512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
DB 6921 TCAGAAATCTCCACAGAGTGCTCATCAAGGGGAAACAGCGCGTTTGTAGTAGGGG 6980
QY 532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
DB 6981 AAAACACATTCAGCTGCTAGTACCAATATGGTGGTGAAGACCAAGAAATGGTCAGAAC 7040
QY 552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
DB 7041 ATACATGTACTCAGCGCAGAGATCAATGCCAGATGGAAACCGAAACTCTGTTCTGAGT 7100
QY 572 HisSerAlaLys----- 575
DB 7101 CACTCGGCAAGGTACGAAATTTTCTGATCATGAGGAATTTTCTTTTAAATTGACTG 7160
QY 575 ----- 575
DB 7161 AATCAACATTTATCTGTATGAAGGAATAATATGGTGCTATCAACATGTTAAGAAATAGC 7220
QY 575 ----- 575
DB 7221 ATACAAATGTTTATTATATGCTTTCCACTGTTCTTCTTACTTATGTTTTCATCTCTT 7280
QY 575 ----- 575
DB 7281 TTGTGTGTGCGTGATGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 7340
QY 575 ----- 575
DB 7341 ACGTGGTGGCGCAATATCTTTTACATCTCATATATATATATATATATATATATATATAT 7400
QY 576 -----ValSerProAlaGluHisAspIleGlnIleMetSerAs 588

DB 7401 TTTTCTCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATTCCAATATATCTCTGA 7460
QY 588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgG1 608
DB 7461 CCTTCATGACGAGTCTACCCCAAGAAAGAAAGCAAAACCTTGAAGTACTGCTGTGA 7520
QY 608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysSerG1 628
DB 7521 AAAACAGACCATATAGATGACATCCCATGATATTTGTGAATCTCTAGCTATAAAACCA 7580
QY 628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLys 648
DB 7581 GCATGAGAGCGAGCTTATGACTGAGACTGATTTGTTCTGACATCAACCGTATTCATCCAA 7640
QY 648 sThrThrAlaAspAspAspCysValIleValAlaLysAspGlySerAspTyrAlase 668
DB 7641 GCACTCTGCTGATGATGATTTGTAATAGTGTGCCAAGGATGGTTTCAGATATATGATC 7700
QY 668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysG1 688
DB 7701 AAGTGTGTTTGACACTAATTCCTCAACAGAGTCTCTGGCATCCCAAGTACAGAAOCA 7760
QY 688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnse 708
DB 7761 GTTACAGGGTCTATTTGGCATTGACCAACAGAGTCTCCACATCTCCAGAACTTTTCAGTC 7820
QY 708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
DB 7821 TACTCAGGAACAGCAGACACATTTGGGATGGAGAAATGGTCACTATTTGCTGCAAGCTC 7880
QY 728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG1 748
DB 7881 ACCACTATTTTCACATCATGATGATCAGTATATTTGCTGAAGCACCACCACTGAACATTTGGG 7940
QY 748 yArgLysAspAlaLysLeuLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
DB 7941 CCGTAAGGACGGAAGAGCTAACCGTGGGAGCAATTTAAGGCCACTACAGAAATTTCTCC 8000
QY 768 oAlaAlaThrCysGlyValaGlnPheArgProGlyIleGlnAlaValaAspLeuThrSerTh 788
DB 8001 AGCAGCAACATGTGGTCTCAATTTAGACTGTGTATCTCAAGCAGATGACTTGACTTCTAC 8060
QY 788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
DB 8061 TCATGTTCATGGGATCTTCAGCAATTTATGATCTGCCAACCAAGTAATTTGCGCATGGA 8120
QY 808 pATsTyrAlaGluArgAlaValaAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
DB 8121 CCGCTATGCTGAAGAGCGGTTAACCCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 8180
QY 828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
DB 8181 AACCATGGAAGCGGTAAGTTATGTGATCGGAAATGTGGACAAGTAGTCTTGTATATCC 8240
QY 848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
DB 8241 TAAAGAAATCCATGCTTGGAGCGATCTTCTGAGATGATGATGATCCATCAACATTAGCAAG 8300
QY 868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG1 888
DB 8301 CTTTCCCAACTATGGAATCTTACGCAAGAACCAAGATGGAGTCTCAACTTCATATATTTCTCA 8360
QY 888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerThrGlySerAsnLeuAsnG1 908
DB 8361 GTATGACATTAATCAGTACAAAGANTCAACCAACATCATATGGCAGTAACTGATATGG 8420
QY 908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
DB 8421 AAAGATTCATTCAGATTCGAGACTTATCAGGCGATTCAGCTGCATGATCTGCACAGACC 8480
QY 928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948

Db 8481 TTACGCCACATCTTAGAGTTGGTGGCTCTCTCTCGCAGAGGAATTCGAAA 8540
Qy 948 nTTPSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleth 968
Db 8541 CTGTCGAGAACTGTGGCACAACTCTGTTATAGTTAGGAGTGTCAACAGGAATAAC 8600
Qy 968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
Db 8601 ATCGATCAGATGAACAGAAAGGAACATTTTGAGCCCTCGAATCTGGAAATGTTTCAGC 8660
Qy 988 aLysTTPAsnAlaLeuGlnLeuGlySerValSerSerAlaAspPheLeuSerAlaAr 1008
Db 8661 AAAATGGAATGATTCAGTTGGTCTCTGTAGCTCCAGTGCAGATTTTATCAGCGAG 8720
Qy 1008 gAnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
Db 8721 GAACAGCATAGCTCAATCTTTGGACAGAGGCAAGGGTAAATGTTCAATCCCTTGGATCG 8780
Qy 1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrIleSe 1048
Db 8781 GTTGTGAGACAGATATCTGTATTAACTAACAGACCCAGCTGATTTTACTACATCAG 8840
Qy 1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
Db 8841 TAACGATAACGAGTATATGGAATACCGC 8868
RESULT 4
ABZ42128
ID ABZ42128 standard; cDNA; 3291 BP.
XX AC ABZ42128;
XX DT 27-FEB-2003 (first entry)
XX DE Arabidopsis thaliana gene #112 modulated by PTGS.
XX EE Posttranscriptional gene silencing; PTGS; plant; transformation; gene;
XX FF ss.
XX GS Arabidopsis thaliana.
XX HH Key Location/Qualifiers
XX IDS 1..3291
XX FT /*tag= a
XX GN WO200281695-A2.
XX PD 17-OCT-2002.
XX PE 05-APR-2002; 2002WO-EP003806.
XX PF 06-APR-2001; 2001US-0282049P.
XX PG (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PH (FRIE-) FRIEDRICH MIESCHER INST.
XX PI Zhu T, Glazov EA, Meins F, Wang X, Chang H;
XX PP WPI; 2003-103337/09.
XX PR P-PSDB; ABP81284.
XX PT Novel polynucleic acid segment useful for modulating gene expression
XX PT within a cell by posttranscriptional gene silencing, and for augmenting a
XX PT plant cell genome.
XX PS Claim 18; Page 403-405; 438pp; English.
XX CC The invention relates to a novel isolated polynucleic acid segment
XX CC modulated within a cell by posttranscriptional gene silencing (PTGS). The
XX CC invention specifically relates to a method to identify an expression
XX CC product that is modulated by PTGS. The polynucleotide is useful for
XX CC modulating the gene expression within a cell by PTGS, by introducing the
XX CC polynucleic acid into a cell and expressing the nucleic acid segment in

CC the cell to form a product. The polynucleic acid segment is also useful
CC for augmenting a cell genome, and for augmenting a plant genome, by
CC contacting a plant cell with the segment to produce a transformed plant
CC cell, and growing the transformed plant cell to produce a differentiated
CC transformed plant. The sequences shown in ABZ42017 - ABZ42142 represent
CC segments of A. thaliana cDNA modulated by PTGS
XX
SQ Sequence 3291 BP; 1066 A; 689 C; 727 G; 809 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.53e-19 Length: 3291
Score: 393.00 Matches: 248
Percent Similarity: 36.81% Conservative: 191
Best Local Similarity: 20.81% Mismatches: 401
Query Match: 7.11% Indels: 352
DB: 7 Gaps: 58

US-09-721-114-2 (1-1057) x ABZ42128 (1-3291)

Qy 55 GlnCysGluHisPheSerIleArgGlyTyrValAlaLeuGlnLysLysAspProLys 74
Db 73 AAATGTGATCATTTTTCATGCTGCTGATTCGTAAGCTCGTGAAGACCTTGA 132
Qy 75 PheCys-----SerLeuSerArgIlePheHisAspGlnLysLysCys 88
Db 133 AAATGTGCTCCGCTTTTCAGAGAGAGTGTAGTTAGTA----- 171
Qy 89 AspGluHisLysAlaSerSerProPheSerValAlaLysPheArgArgTirPaspCys 108
Db 172 GACCAACAAAGCTATCTTCTACTTATCTTCTTCCAAAGCTTTAGATGGCATCT 231
Qy 109 SerLysCysLeuAspLysLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
Db 232 ATGAGCTGCATCAAGATATA-----GATGCTCATGGGCCAAAGATTTGTGA 279
Qy 129 AlaLysGlnAsnGlyThrSerAspGly---CysSerIleThrPheValArgSerThrPhe 147
Db 280 CTGCATCAACTCAAAAGCTATTCGAAACTCTTCTGTTATCGAAAGTAAAGCAAGTTC 339
Qy 148 ValProAlaSerValGlySerGlnLysValSerProSerThr----- 161
Db 340 AATTCGTAACATCATTCATCAGAGAAAGAAAGAAACTGATATTGCAGATAATGCT 399
Qy 162 ---GlnSerSerGlnGlyLysAsnAlaAspArg-----SerThrLeuPro 175
Db 400 ATTGAGGAGAAAGTGGGTAAACTGTGAGATGATGATGATGATGATGATGATGAT 459
Qy 176 LysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAla 195
Db 460 AAGAAAGCACGT-----GGTCGACCTATCGGT----- 486
Qy 196 AlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAsp 215
Db 487 ---GCTTCTTAATGTTAGGAGCAAGAGCAAGCTGTGAGTCCGAGAGC----- 534
Qy 216 ValAlaAlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeu 232
Db 535 GTAGGAACACACAGATCTAAGAAACAACTAAACAAACCATCGATGATATTAGCAGCTG 594
Qy 233 ProGluValProGlnIleThrTrpHisLeuGluValAsnGlyAla----- 247
Db 595 AAAGAGAAACAAATGTGATCAGCTGTGACAAAGCTTGGCTCATCTCAAAATTCCTGT 654
Qy 248 -----AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsn 264
Db 655 GTGTTTGGAGTACACACCTTAAGCAACCAAGATCATATAAGGCAATTCGGCTGTGATG 714
Qy 265 GluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLys 284
Db 715 GAATGCGATAACGGGTGCATCAGAAAGTATAAATCTTGT----- 753
Qy 285 AspProAsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCys 301

Qy	592	GlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThr	611
Db	1759	CATGGAGCACCAAGAGCAAGACTCAGATAGAAAGGAGACTACGACTGAAGGCAAAAC	1818
Qy	612	MetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArg	631
Db	1819	AAC--GATGATATTCCAATGGAGATAGTGGAGCTCATGGCCAAAACCACTACGAGAGG	1875
Qy	632	GlnLeuMet---ThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysThr	650
Db	1876	TGCTCTTCCGACAAAGAGAAGATGTTAGCAACAAACAGCCATCAACAAGAAACAGCACAC	1935
Qy	651	AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal	670
Db	1936	AAATCCAGAATGCTCTACTGATTGATCTCAATGAAACCTACGATAACGGGATCTCACTT	1995
Qy	671	PheAspThrAsnSerGlnGln-----LysSerLeuAlaSerGlnSerThrGlnLysGlu	688
Db	1996	GAGGACACCAACACATCAGACCAACAAACCGGTAGTAGCAACGCAAGAGGGAGAA	2055
Qy	689	LeuGlnGlyHisLeuAlaLeuThrGlnGluSerProHis-----	702
Db	2056	-----CAFTTCTCTANGGAAGACAGCACACTCTCATGACTTCTTCCCAATAAGT	2106
Qy	703	-----ProGlnAsn-----PheGlnSerThrGlnGluGlnGlnThrHis	715
Db	2107	CAGCCTTATGTCGCTTCTCGTTTGGATGCTTCTTCCTACCCCAAGAAACCCGA-----	2160
Qy	716	LeuArgMetGluGluMetValThrIleAlaAspSerProLeuPheSerHisAsp	735
Db	2161	-----GCCAGCTCCATCCGGTTTCTTGGTGCACAAC	2190
Qy	736	AspGlnTyrIleAlaGluAlaProThrGluHisTyrGlyArgLysAspAlaLysLysLeu	755
Db	2191	TGTCATGGCTCGGGAATTGGCAACTGTG-----GGTAATCAGAACCCCTTCTCCATCC	2244
Qy	756	ThrTyrGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAla---	774
Db	2245	TCATTTCCGGTATTACGTGCT-----TGTAATCTTGCCAGAGTGT	2286
Qy	775	-----GlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGly	792
Db	2287	CTTAATCAATACAGA-----GAGCTTCTCATCCAAITGG	2322
Qy	793	SerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaGlu	812
Db	2323	CCATCTTCC-----ATGATACCACACACAGAGTCAATACAAGCCA	2361
Qy	813	ArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAla	832
Db	2362	GTTCCTTTAAATATTAAATCAGTCACAATAT---CCGGTACCGCTTTCA-----CAGGCA	2412
Qy	833	SerLysLeuCysAspArgArgAlaGlyGlnValValLeuTyrProLysGlu-----	850
Db	2413	TCTAACCAATGAAATATACATGGAACTCAACTTTGTTGCTGCCAACGGGAAGCAAAATGT	2472
Qy	850	-----	850
Db	2473	GGGCCTAATCCAGAAATTTCATTGGCTGCMAACATGCTGGGGTTAGTAGTAGTAGT	2532
Qy	851	-----SerMetProAlaThrHisLysLeuArg	859
Db	2533	AGTAGGCCAATAGATAACTTTTCTAGTAGAGCTCTATACCGGCATTTGCTACTCTACG	2592
Qy	860	MetMetAspProSerThrLeuAlaSerPhePro-----AsnTyrGlyThrSer	875
Db	2593	CTTCTGATCCTCGCTGAGTCAACACTCCGCTGACCAACGGAACACTAAATTT	2652
Qy	876	SerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLys	895
Db	2653	ACTAAAGACATTTTCGGCAGCCACACAGTCCAAAGAGTTTATAGAGTCTCAACAGGG	2712

transgenic plant; flowering control; reproductive trait control.	Arabidopsis thaliana.	Location/Qualifiers
Key	1..3201	
Promoter	/tag= a	
exon	3202..3265	
	/tag= c	
intron	/number= 1	
	3266..4159	
	/tag= d	
	/number= 1	
exon	4160..4335	
	/tag= e	
	/number= 2	
CDS	4241..8031	
	/tag= b	
intron	/product= "Arabidopsis thaliana EMP1 protein"	
	4336..4447	
	/tag= f	
	/number= 2	
exon	4448..4623	
	/tag= g	
	/number= 3	
intron	4624..4703	
	/tag= h	
	/number= 3	
exon	4704..4823	
	/tag= i	
	/number= 4	
intron	4824..4902	
	/tag= j	
	/number= 4	
exon	4903..4956	
	/tag= k	
	/number= 5	
intron	4957..5045	
	/tag= l	
	/number= 5	
exon	5046..6307	
	/tag= m	
	/number= 6	
intron	6308..6447	
	/tag= n	
	/number= 6	
exon	6448..8065	
	/tag= o	
	/number= 7	
intron	8066..8300	
	/tag= p	
	/number= 7	
exon	8301..8648	
	/tag= q	
	/number= 8	
US3676751-B1.		
23-APR-2002.		
08-OCT-1999;	9US-00415946.	
09-OCT-1998;	9US-00169696.	
(REG) UNIV CALIFORNIA.		
Sung ZR, Aubert D, Chen L;		
WPI: 2002-453154/48.		
P-PSDB; AAO15030.		
A polynucleotide which hybridizes to an embryonic flower 1 probe promotes early transition from vegetative to reproductive state		

Arabidopsis thaliana embryonic flower 1 (EMF1) genomic sequence.

A polynucleotide which hybridizes to an embryonic flower 1 protein and promotes early transition from vegetative to reproductive state when in a

plant is useful to control flowering in transgenic plants.

Claim 3; Col 35-48; 30pp; English.

The invention comprises the amino acid and coding sequence of the Arabidopsis thaliana embryonic flower 1 (EMF1) protein. The EMF1 nucleotide of the invention promotes early transition from a vegetative to a reproductive state when operably linked to a plant promoter and introduced into a plant. The EMF1 nucleotide promotes early transition to a reproductive state by hybridizing to the EMF1 gene that is naturally present within the plant. The EMF1 DNA sequence is useful for controlling flowering and other reproductive traits in plants. The present DNA sequence represents the Arabidopsis thaliana embryonic flower 1 (EMF1) genomic sequence

Sequence 8648 BP; 2748 A; 1563 C; 1631 G; 2706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,11e-11	Length:	8648
Score:	298.00	Matches:	225
Percent Similarity:	35.53%	Conservative:	152
Best Local Similarity:	21.21%	Mismatches:	360
Query Match:	5.39%	Indels:	324
DB:	6	Gaps:	50

US-09-721-114-2 (1-1057) x AAL43177 (1-8648)

Qy	198	AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValala	217
Db	5090	TCTAATGTAGAGCAAGAGCAGAGAGCTTGAGTCCGAGCAG-----GTAGGA	5140
Qy	218	AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu	234
Db	5141	AACACACAGATCTAAGAGAAAACCTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGG	5200
Qy	235	ValProGlnIleThrTPhisIleGluValAsnGlyAla-----	247
Db	5201	AAACAAATGGGATCGGCTGTGACACAGCTTCGGCTCATCTGAAATGGTGGTGGT	5260
Qy	248	---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp	266
Db	5261	GAGGATACACCCATCAAGCAACCAAGCAATCAATAGGCATTCGCGGTCTGTGGAATGC	5320
Qy	267	GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro	286
Db	5321	GATACGGGTCTACAGAAAGATTAATCTTGCT-----	5353
Qy	287	AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu	303
Db	5354	-----ATGAGTGGGTTCAGCGGTAGGAAATCTCCCAAGGTTCGTCTACTCAGTGAGTTG	5407
Qy	304	ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys	323
Db	5408	CTTGGTAATACAAAACCAAGCGGTGGTAGTAACATCAGAAAA-----	5449
Qy	324	GluGluValValLeuLysArgSerSerLysArgLysThrAspLysLysLeuMet	343
Db	5450	GAAGAGTCTGCTTGAAGAGGATCAGTTAGAGGTGMAAA-----AGAAAGTGTGA	5503
Qy	344	LysLysGlnGlnHis-----SerLysLysArgThrAlaGlnAlaAspValSerAsp	360
Db	5504	CCTGAAACAAATATTATGACCGCGATATTGAGTCAATGGGTGCAACCTCTGAAAAATGCT	5563
Qy	361	AlaLysLeuCys-----	364
Db	5564	TCCAAAGTGTGACTCTGATCAAGGTAATAGTCAATCACTGATAGTGGTTGACAGA	5623
Qy	365	-----ArgArgLysProLysValArgLysValArgLeuLeuSerGluIleAsn	379
Db	5624	ACTCCATTAAAGCGGTAAAGCAGAGAAACAGAGATTTCAGGTGTGACGAGTTGTACCA	5683
Qy	380	AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAspPro	399
Db		-----	

Db	5684	TCACCTTCTGTGAAACTTCACAA-----GAAGGTATCAAGGAGCATGATGATGATCCT	5737
Qy	400	CysGluAspAspArgSerThr-----	408
Db	5738	AGTAAG-----AGATCAACTCTCTGCGACTCTTTATTCTACTCGAAACGATTCGTCTCT	5791
Qy	409	ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyLysAspGly	428
Db	5792	TGTCT-----COGGTACTCAGAGAAC-----CAGAGGAAG	5824
Qy	429	LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer	448
Db	5825	CTCAGTTTATCCCAAGAGAACAGACAAAG-----CCTGTATTCGATATATGGAAG	5875
Qy	449	SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly	462
Db	5876	AGCACTGTGATCAGTTTACTAAGCGCATTTAGTGAAGTCAAGTAACTGCGATACCTGCT	5935
Qy	463	SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro	482
Db	5936	CTTTCCATGAACACAGTATCCCAACTCGAGACTTATTGATCGGAAGGGTGGGGCGGT	5995
Qy	483	ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp-----	499
Db	5996	TTATTTGACAAACCGTTTGGCTTCAGATGGATATTTTCAGAAATATCTCTCTCAGGTTAAT	6055
Qy	500	-----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer	515
Db	6056	GATAGCCGATACATCTTTGCAATTTGCAAGACAATGATTTATCTGAGGTCAAGACGCG	6115
Qy	516	ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly-----Leu	528
Db	6116	GAACCAAACTGCTTCGAGATTTTAGTTCTCTCTTAAATCCAGCTCAGGTGGATGGTTG	6175
Qy	529	SerLysGly-----	533
Db	6176	AGAATCGGAGTAGATATTTGACTTTCAGAAACCAACCAACATACAAACAGATGTCT	6235
Qy	534	HisSerAlaAlaSerThrLysThrGlyGluSerThrArgAsnGlyGlnAsnIleHis	553
Db	6236	TTCTCGAACTTGAAGCTAAGATACCCCTCTCTCTACGAGGTCCGGAATTTATCTCGG	6295
Qy	554	ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe	573
Db	6296	GTGCTCAAAAGGTACAGCTTTGCATCTATTTCTTCAATCGCCATCTCTTCTCATTC	6355
Qy	573	-----	573
Db	6356	TGTTTTCCAACCTGACCTATTCTCTGTATATTCCAAACCTGGTGTCTTCAGTGATGGAT	6415
Qy	574	-----Ala-LysValSerProAlaGluHisAspIleG	584
Db	6416	CAGGATGATAACGGGATTGATTTACTTGGCAGGATCTTCTGTCAGATAGAAAGGGA	6475
Qy	584	InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG	604
Db	6476	AGACTGTTATGTGTCAGAACATCATATGAGCACCAAGAACCAAGCTCAGATAGAAAGG	6535
Qy	604	IuValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL	624
Db	6536	AGACTACGACTGAAGAGCAAAACAAAC---GATGATATTCATATGGAGATAGTGAGCTCA	6592
Qy	624	euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA	643
Db	6593	TGSCCAAAACCAAGTACGAGAGGTGCTTCCGACAAAGAGAGATGTTAGCAACAAAC	6652
Qy	643	snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaLysAspG	663
Db	6653	AGCCATCACAGAAACAGACACAAATCCAAAGATGCTCTACTGATTTGATCTCAATGAAA	6712
Qy	663	lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA	681
Db	6713	CCTACGATTAACGGATCTCATTGAGACAAACACACATCAGACCACCAAAACCGTGTA	6772

```
QY 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 6773 GTAGCAACCGAAGAGGAGAA-----CATTTTCTATGGAAGACAGCAACT 6823
QY 701 roHis-----ProGlnAsn-----PheGlns 708
Db 6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTCCTCTCCGTTTGGCATCTTCTCTC 6883
QY 708 erThrGlnGlnGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db 6884 CTACCCCAAGAAACCGA-----GCCAGCT 6907
QY 728 exProLeuPheSerHisAspAspGlnTyrlleAlaGluAlaProThrGluHisTrpG 748
Db 6908 CCATCCGGTTTCTGGTCACAACCTGTCAGNGCTTGGGAATTTGCCAAGCTGTG-----G 6961
QY 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 6962 GTAATCAGAACCCCTTCTCCATCCTCATTTCCGGTATTACGTCT----- 7005
QY 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Db 7006 --TGTGATCTTGGCAGAGTCTTCTTAATCATACAGA----- 7041
QY 785 euThrSerThrHisValMetGlySerSerSerAsnTyrlaSerArgGlnProValIleA 805
Db 7042 --GAAGCTTCTCATCCCAATTTGGCCATCTTCC-----ATGATAC 7078
QY 805 laProLeuAspArgTyrlaGluAlaValHisGlnValHisAlaArgAsnPheProS 825
Db 7079 CACCACAGAGTCATACAGCAGCTTCTTTAATATATATCATCTCAACAAT-----CCGG 7135
QY 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db 7136 GTAGCGTITTC-----CAGCACTCAACAATGAATACATGGAACCTCAACTTCTGTG 7189
QY 845 alLeuTyrlProLysGlu----- 850
Db 7190 CTGCCAAGCGGAGCAAAATATCGCGCTAATCCAGAAATTTTCATTTGGTGTCAAAACATG 7249
QY 851 -----SerM 852
Db 7250 CTGCTGGGTTAGTAGTAGTAGTAGTAGGCGCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309
QY 852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro----- 870
Db 7310 TACCGCATTCATCTACTCAGCCTTCTGGATCCTCGCTGAGGTCAACGACTCCCGCTG 7369
QY 871 --AsnTyrlGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 7370 ACCAACCGGAACACTAAATTTACTTAAAGACATTTCCGCCAGCCAAACGATCCCAAG 7429
QY 888 lntyrlAlaHisAsnGlnTyrlLysGlySerThrSerThrSerThrGlySerAsnLeuAsnG 908
Db 7430 AGTTTATAGAGCTTCAACACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG----- 7482
QY 908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db 7483 --CAGATACCT-----TTTGATCTATACAGAGAGATTACACAGAGCCTTCCCGGA 7534
QY 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db 7535 AGAGTTTCCCATCACTCCACTATTTGGAGC-----TCTTCACTTTCATTTCAAAATG 7588
QY 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrlLysLeuGlyValSerThrGlyI 967
Db 7589 CTTTCATGGAGTCTCAT----- 7605
QY 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 7606 -----CATCAGAGAGAGAAACCAAGAGAAAGACACC-----TTTG 7642
```

```
QY 987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 7643 CTCCTGTTTACAACTACTCATGAAAGACCGGTGTTTTCGAAGCAGCAACGACCAAGCGAAGT 7702
QY 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db 7703 TCCAGCTCTGGAGACATCGAATTC-----A 7729
QY 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 7730 TGATGCTTCTTGAATTTTCAATGACGATATAAGAAAGAAACAAAGAGAAAGAGCAG 7789
QY 1034 ----- 1035
Db 7790 AGAGCTGCAATACCAATGCTCTGCGGACCTGTGAAGAACAGTCTTGACCCATTTGTT 7849
QY 1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGlnTyrlMet 1054
Db 7850 GCAGCGTCATAGAAACCTCTGTGATTCACCATTCCTGACCTGGGAATGTTTACATG 7908

RESULT 7
ABQ777909
ID ABQ777909 standard; DNA; 8648 BP.
XX AC ABQ777909;
XX DT 24-JAN-2003 (first entry)
XX DE Arabidopsis thaliana EMFL gene.
XX EMFL; embryonic flower 1; rice OsEMFL1 homologue; plant; floral repressor;
KW reproductive development; flower development; transgenic plant;
KW antisense suppression; transgenic; reciprocal negative interaction;
KW flower meristem identity gene; flowering time; shoot development;
KW seed yield; agriculture; gene; ds.
XX OS Arabidopsis thaliana.
XX PH Location/Qualifiers
FT Key 4241..8031
FT CDS /*tag= a "EMFL protein"
FT FT 4241..4335
FT FT /*tag= b
FT FT /number= 1
FT FT 4335..4447
FT FT /*tag= c
FT FT /number= 1
FT FT 4448..4623
FT FT /*tag= d
FT FT /number= 2
FT FT 4624..4703
FT FT /*tag= e
FT FT /number= 2
FT FT 4704..4823
FT FT /*tag= f
FT FT /number= 3
FT FT 4824..4902
FT FT /*tag= g
FT FT /number= 3
FT FT 4903..4956
FT FT /*tag= h
FT FT /number= 4
FT FT 4957..5045
FT FT /*tag= i
FT FT /number= 4
FT FT 5046..6307
FT FT /*tag= j
FT FT /number= 5
FT FT 6308..6447
FT FT /*tag= k
FT FT /number= 5
FT FT 6448..8031
FT FT exon
```


QY 554 ValIeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValIeuSerHisSe 573
Db 6296 GTGCTGCAAAAGGTACAAAGCTTTGCTATCTATTTCTTCAATGCCAGCTTCTCAATC 6355
QY 573 r----- 573
Db 6356 TGTTTTCCAACTGCACCTATTCTTGTATATTCCAAACCTGGTGTCTCAGTGATTCGAT 6415
QY 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
Db 6416 CAGGATGATACCGGATTTGATTTACTTGGCAGGATGCTTCTGTCAGATAGAAAGGGA 6475
QY 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG 604
Db 6476 AGACTGTTATGGTCCAGAACATCATCGAGCACCAAGCAAGCAAGTCAAGATAGAAAGG 6535
QY 604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 6536 AGACTTACGACTGAAGCAACAAACAC---GATGATATTTCCAAATGAGATAGTGGAGTCA 6592
QY 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 6593 TGGCCAAACCACTACGAGAGGTCTTCCGACAAAGAGAGATGTTAGCAACAAC 6652
QY 643 snArgIleGlnSerLysThrAlaAspAspAspCysValIleValAlaLysAspG 663
Db 6653 AGCCATCACAGAAACACGACACACAAATCCAAAGATGCTTACTGATTCATCAATGAA 6712
QY 663 LysSerAspTyAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA 681
Db 6713 CCTAGATAACGGATCTCCTTACGAGCAACACATCATCAAGACCAACCAACCGGTGTA 6772
QY 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnLysuSerP 701
Db 6773 GTAGCAACGCAAGGAGGCAAGAA-----CATTTCTCTATGGAAGACAGCAGAACT 6823
QY 701 rOHis-----ProGlnAsn-----PheGlns 708
Db 6824 CTCATGACTTCTCCCAATAAGTCAGCCTTATGTCCTTCTCCCTTGGGATCTTCTCTC 6883
QY 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaLysSerS 728
Db 6884 CTACCCAGAAACCGA-----GCCAGCT 6907
QY 728 erProLeuPheSerHisAspAspGlnTyrlIleAlaGluAlaProThrGluHisTrpG 748
Db 6908 CCATCCGGTCTTCTGTCACAACTGTCAGTGGCTTGGGAATTTCCCAACTGTG-----G 6961
QY 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 6962 GTAATCAGACCTTCTCCATCTCTCATTTCCGGTATTAGTGCT----- 7005
QY 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
Db 7006 --TGTGATACTTGCAGAGTGTCTTAATCAATACAGA----- 7041
QY 785 euThrSerThrHisValMetGlySerSerSerAsnTyrlAlaSerArgGlnProValIleA 805
Db 7042 --GAAGCTTCTCAATCCAAATTTGGCCACTTCC-----ATGATAC 7078
QY 805 laProLeuAspArgTyrlAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db 7079 CACCACAGTCAATACAGCCAGTTTCTTTAAATATTAAATCAGTCAACAAT---CCGG 7135
QY 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db 7136 GTACGCTTTCA-----CAGGCATCTAAACAAATGAAATATACATGGAACCTCACTTTGTG 7189
QY 845 alLeuTyProLysGlu----- 850
Db 7190 CTGCCAAGGAGCAAAAATGCTGGCCTAATCCAGAAATTTTCATTTTGGCTGCAACATG 7249

851 -----SerM 852
Db 7250 CTGCTGGGTTAGTAGTAGTAGTAGCCCAATACATTTTCTAGTGAGAGCTCTA 7309
QY 852 euProIaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro----- 870
Db 7310 TACCGCATTTGATCTACTACTCAGCTTCTGGATCTCTGCTGAGGTCAACAGACTCCCGCTG 7369
QY 871 --AsnTyrlGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 7370 ACCAACACGGAAACACTAATTTACTAAAGACATTTTCCGCCAGCAACCACTCCCAAG 7429
QY 888 InTyrlAlaHisAsnGlnTyrlLysGlySerThrSerThrSerThrSerThrSerThrSer 908
Db 7430 AGTTTATAGAGCTTCAACACAGGGGACTCTAGTAAAGTCAGCTACTCAACTAAG----- 7482
QY 908 LysLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db 7483 --CAGATACT-----TTTGATCTATCAGCAAGATTCACACAGAGCTTCCCGGA 7534
QY 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db 7535 AGAGTTTCCCATCTACTCCACCTATTGGGACG-----TCTTCACTTTCACTTCAAAATG 7588
QY 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrlLysLeuGlyValSerThrGlyI 967
Db 7589 CTTATGGAGTCTCTCAT----- 7605
QY 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db 7606 -----CATCAAGAGAAAGAAACCAAGAGAAAGACACC-----TTTG 7642
QY 987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
Db 7643 CTCTCTTTTACATACTCATGAAAGCCGGTGTGTTGCAAGCAGCAACGACCAAGCGAAGT 7702
QY 1004 ----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db 7703 TCCAGCTGTGGGAGCATCGAATTC-----A 7729
QY 1022 etValHisProLeu-----AspArgPheValArgGlnAsp----- 1033
Db 7730 TGAATGCTTCTTTGAAATTTTCATGACGGATAAAGAAAGAAAGAGAGAAAGCAG 7789
QY 1034 -----IleC 1035
Db 7790 AGAGCTGCAATACAAATGCTCTGCGGACCTGTGAGAACAGCTTCTGGACCCATTGTG 7849
QY 1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrlMet 1054
Db 7850 GCAGCGTCAATAGAAACCTGCTGATTTCCCACTTCTCGAACCTGGGAATGTTTACATG 7908

RESULT 8

ANAL4872
ID AAA14872 standard; DNA; 17341 BP.

XX AC AAA14872;

XX DX 08-AUG-2000 (first entry)

XX DB Genomic DNA sequence from Arabidopsis thaliana CD82 clone.

XX XX Embryonic flower gene-1; EMF-1; reproductive development; flowering;

XX KW early flowering; uniform flowering; ss.

XX XX Arabidopsis thaliana.

XX OS Arabidopsis thaliana.

XX XX WO200022132-A2.

XX PN 20-APR-2000.

XX PD 08-OCT-1999; 99WO-US023543.

PR 09-OCT-1998; 98US-00169696.
 XX (REGC) UNIV CALIFORNIA.
 XX PI Sung ZR, Aubert D, Chen L;
 XX WPI; 2000-317987/27.
 DR An isolated nucleic acid molecule comprising embryonic flower genes is
 XX useful for modulating reproductive development in plants.
 XX Claim 3; Page 30-33; 38pp; English.
 XX The present sequence represents the genomic sequence of the Arabidopsis
 CC thaliana genomic DNA from CD82 clone. The sequence comprises the
 CC embryonic flower gene-1 (EMF-1). The EMF genes play an important role
 CC in reproductive development in plants. Control of the expression of the
 CC genes is therefore useful in controlling flowering and other functions in
 CC plants. The EMF-1 polynucleotides are useful for modulating reproductive
 CC development in plants. For example, enhanced expression of EMF-1 is
 CC useful to prevent flowering in grasses or vegetable crops. Controlling or
 CC inhibiting expression of the genes is useful to promote early flowering
 CC in forest trees and agronomical crops or to ensure uniform flowering in
 CC various agronomic crops or ornamentals.
 XX
 SQ Sequence 17341 BP; 5468 A; 3114 C; 3409 G; 5350 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,78e-11 Length: 17341
 Score: 298.00 Matches: 225
 Percent Similarity: 35.53% Conservative: 152
 Best Local Similarity: 21.21% Mismatches: 360
 Query Match: 5.39% Indels: 324
 DB: 3 Gaps: 50

us-09-721-114-2 (1-1057) x AAAL4872 (1-17341)
 QY 198 AlaAsnThrAspSerProMetIysAspLeuGlnGlyProAlaGlnAsnTyrAspValala 217
 DB TCTATGTTAGGACGACGAGGAAAGCTTGTGAGTCCGGAGCG-----GTAGGA 5140
 QY 218 AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu 234
 DB TCTATGTTAGGACGACGAGGAAAGCTTGTGAGTCCGGAGCG-----GTAGGA 5140
 QY 235 ValProGlnIleThrPheIleGluValAsnGlyAla----- 247
 DB 5201 AAACAAATGTGGATCAGGCTGTGACACAGTTCGGCTCATCTGAAATGTGGTGTGTT 5260
 QY 248 ---AspGlnProProSerThrProIysLeuSerGluValValLeuLysAsnGluAsp 266
 DB GAGGATACACACCTAAGCAACCAAGAAATCAATAAGGCAATTCGGGTCGATGGAATGC 5320
 QY 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
 DB GATAACGGGTCTACAGAAAGTATAAATCTGCT----- 5353
 QY 287 AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu 303
 DB ATGAGTGGGTTCGACGCTAGGAAATCTCGCAGGTTCTCTACTCAGTGAGTTG 5407
 QY 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
 DB CTGTGTAATACAAACCAAGGCTGTGAGTACATCAGAAAA----- 5449
 QY 324 GluGluValValLeuValArgSerLysSerLysAspValSerLysThrAspLysValLeuMet 343
 DB GAAGAGTCTGCTTTGAAGAGGAATCAGTTAGAGTGTGAAAA-----AGAAGTTGTTA 5503
 QY 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp----- 360
 DB CCTGAAACAAATATGTCTACGCGGATATTGAGTACATAGGCGTCAACCTCTGAAATGCT 5563
 QY 361 AlaLysLeuCys----- 364
 DB TCCAAAGTTGTGACTCTGTGATCAAGGTAAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623
 QY 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluIleAsn 379
 DB ACTCCATTAAAGGTAGCAGAGAAACAGAAAGTTTCGGTTGTGTGACGAGTTGTGTACCA 5683
 QY 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
 DB TCACTTCCTTGTGAACACTTCACAA-----GAAGGTATCAAGGAGCATGATGACAGATCCT 5737
 QY 400 CysGluAspAspArgSerThr-----IlePro 408
 DB AGTAAG-----AGATCAACTCTCGGCACCTCTTTATTCATCTGGAACAGATCTGTCTCT 5791
 QY 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
 DB TGTCT-----CGGGTACTCAGAGAAC-----GAGAGGAAG 5824
 QY 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
 DB CTTCAGTTTACCAAGAGAAAGACAAAGAG-----CCTGTATCGATATATGGAAG 5875
 QY 449 SerLeuMet-----AsnTyrLeuAsnGlyLysLys-----LysArgThrGly 462
 DB AGCACTGTGATCAGTTTGTAGTAAGCGCATTCATGAGAGTCAAGTTAACTCGCATCTGCT 5935
 QY 463 SerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
 DB CCTTCATGAACACACAGTATCCCAACTTCGAGACTTATTTGAATGGGAAAGGTTGGCGGT 5995
 QY 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp----- 499
 DB TTATTTGACAAACCGTTGGCTTCAGATGGATATTTTCAGAAAAATATCTCTCAGGTTAAT 6055
 QY 500 -----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
 DB GATAAGCGGATACATCTTTGCTATTCGAGCAATGATTTATGAGTCAAGAGACGCG 6115
 QY 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu 528
 DB GAACCAAACTGTCTTCGAGATTTTAGTTCTCTCTTCTTAAATCCAGCTCAGGTGGATGTTG 6175
 QY 529 SerLysGly-----LysThr 533
 DB AGAAGTGGATGATATTTGTGATTCAGAAACACCAACCAACATACAAACAGATCTGCT 6235
 QY 534 HisSerAlaAlaSerThrLysTyrGlyGlyGlySerThrArgAsnGlyGlnAsnIleHis 553
 DB TTTCTGAACTTGAAGCTAAGATACCCCTTCTTCTACTGAAAGTTGCGGATTTATCTCGG 6295
 QY 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
 DB GTCTGCAAAAGGTACAGCTTTGCATCTATTTTCTTCATCGCCATGCTTCTCTCATTC 6355
 QY 573 r----- 573
 DB TGTTCACACTTGCACCTATCTCTGTATATTCACAACTGGTGTGCTTCACTGATTTGGAT 6415
 QY 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
 DB CAGGATGATAACCGGATTTGATTTACTTGGCAGGATGCTTCTGTGCTGATAGAAAGGGA 6475
 QY 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysLysLysLys 604
 DB AGACTGTTATGGTCCAAAGACATCATCGAGCAACCAAGAGCCAAAGTCACGATAGAAAGG 6535
 QY 604 LuValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeu 624
 DB AGACTACGACTGAAGAGCAAAACAAAC---GATGATATTCATGAGATGATGAGATCTCA 6592
 QY 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643

plant is useful to control flowering in transgenic plants.

Example 1; Col 19-34; 30pp; English.

The invention comprises the amino acid and coding sequence of the Arabidopsis thaliana embryonic flower 1 (EMF1) protein. The EMF1 nucleotide of the invention promotes early transition from a vegetative to a reproductive state when operably linked to a plant promoter and introduced into a plant. The EMF1 nucleotide promotes early transition to a reproductive state by hybridizing to the EMF1 gene that is naturally present within the plant. The EMF1 DNA sequence is useful for controlling flowering and other reproductive traits in plants. The present DNA sequence represents the Arabidopsis thaliana embryonic flower 1 (EMF1) gene sequence contained within a pBluescript vector

SQ Sequence 17341 BP; 5468 A; 3114 C; 3409 G; 5350 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,78e-11	Length:	17341
Score:	298.00	Matches:	225
Percent Similarity:	35.53%	Conservative:	152
Best Local Similarity:	21.21%	Mismatches:	360
Query Match:	5.39%	Indels:	324
DB:	6	Gaps:	50

US-09-721-114-2 (1-1057) x AAL43176 (1-17341)

Qy	198	AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValala	217
Db	5090	TCTAATGTTAGGCAAGCAGCAGCAAGCTTGTAGTCGGAGCAG-----GTAGGA	5140
Qy	218	AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu	234
Db	5141	AACACAGACTTAAGAAATAACTAAACMAACCATCGATGATATAGCAGCTGGAAGAG	5200
Qy	235	ValProGlnIleThrTrpHisIleGluValAsnGlyVala-----	247
Db	5201	AAACAAATGTGGATCAGGCTGTGCACACGTTGGCTCATCTGAAATTCGCGTGTGTT	5260
Qy	248	---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp	266
Db	5261	GAGGATACACCACTTAAGGCAACCAAGAAATCATAAAGGCATTCGCGTGTGATGGAATGC	5320
Qy	267	GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro	286
Db	5321	GATAAAGGTCATCAGAAAGTATAAATCTGCT-----	5353
Qy	287	AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu	303
Db	5354	-----ATGAGTGGGTTGCAGCTAGGAAATCTGCAAGGTTCTACTCAGTGAGTTG	5407
Qy	304	ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys	323
Db	5408	CTTGTTATACAAAACCAAGTGGTGTAGTACATCAGAAA-----	5449
Qy	324	GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLeuMet	343
Db	5450	GAAGAGTCTGCTTTGAAGAAGGATCAGTTAGAGCTCGAAA-----AGAAAGTTGTTA	5503
Qy	344	LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp-----	360
Db	5504	CCTGAAACAAATATATGTCAGCCGATATTGAGTCAATGGGTGCAACCTCTGAAATGCT	5563
Qy	361	AlaLysLeuCys-----	364
Db	5564	TCCAAAAGTCTGACTCTGTATCAAGGTAATAGTGAATCAACTAGTGGGTTTGACAGA	5623
Qy	365	-----ArgArgLysProLysLysValArgLeuSerGluIleAsn	379
Db	5624	ACTCCATTAAAGGGTAACAGAGAAACAGAGANVTTGAGTTGTTGACGAGTTGTACCA	5683
Qy	380	AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAspPro	399

Db	5684	TCATCTCTTGTGAACACTTCACAA-----GAAGGTATCAAGGAGCATGATGATGATCCT	5737
Qy	400	CysGluAspAspArgSerThr-----IlePro	408
Db	5738	AGTAAG-----AGATCAACTCTCTGCGCATCTTTATTCACTGAAACGATTTCTGTTCT	5791
Qy	409	ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly	428
Db	5792	TGTCCT-----CCGGTACTCAGAGAACA-----GAGAGGAAG	5824
Qy	429	LeuLysSerSerLysAsnLysThrLysArgLysTy-SerAspValValAspAspGlySer	448
Db	5825	CTCAGTTTACCCCAAGAGACAGAGAGAG-----CCTGTATCGATATATGGAAG	5875
Qy	449	SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly	462
Db	5876	AGCACTGTGATCAGTTTATAGTAACGCGCATGTAGTGAAGTCAAGTTAATCTCGCATCTGTT	5935
Qy	463	SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro	482
Db	5936	CCTTCATGACACAGTATCCCAACTCGAGACTTATGATGGGAAAGGGTGGCGGT	5995
Qy	483	ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp-----	499
Db	5996	TTATTTGACAAACCGTTTGGCTTCAGATGGATATTTTCAGAAAATATCTCTCAGGTTAAT	6055
Qy	500	-----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer	515
Db	6056	GATAGCGGATACATCTTTGCATTTTGCAGACATGATATGTGAGGTCAAGAGCGCG	6115
Qy	516	ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu	528
Db	6116	GAACCAACTGCTCTCGAGATTTTGTCTCTCTTCTTAAATCCAGCTCAGGTGATGTTG	6175
Qy	529	SerLysGly-----LysThr	533
Db	6176	AGAAGTGGAGTAGATATTGTGACTTCAGAAACAAACCAACAAATCAACAGATGCTCT	6235
Qy	534	HisSerAlaAlaSerThrLysTyThrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis	553
Db	6236	TTCTCGAAGTCTGAGCTAGATACCCCTCTCTTCTACTAGTTCGGATTTATCTCG	6295
Qy	554	ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe	573
Db	6296	GTCTGCAAAAGGTACAAAGCTTTGTCATCTATTTTCTCAATCGCCATGCTTTCTCATTC	6355
Qy	573	-----	573
Db	6356	TGTTTCCAACTTGCACCTATTCTTGTATATTCCAACTGGTGTCTTCACTAGTATGGAT	6415
Qy	574	-----Ala-LysValSerProAlaGluHisAspIleG	584
Db	6416	CAGGATGATAACGGGATTGATTTTACTTGGCAGGATGCTTCTGTCAGATAGAAAGGGA	6475
Qy	584	IntLeuMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG	604
Db	6476	AGACTTTATGGTCCCAAGAACATCATGGAGCCACAGAAAGCCAAAGTCCAGTAGAAAGG	6535
Qy	604	IuValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL	624
Db	6536	AGACTAGCTGAGAGGCAAAACAAAC---GATGATATTCCATGGAGATAGTGGAGCTCA	6592
Qy	624	euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA	643
Db	6593	TGGCCAAAACCAAGTACGAGAGGTGCTTCCGACAAAGAAAGATGTTAGCAACAAAC	6652
Qy	643	snArgIleGlnSerLysThrThrAlaAspAspCysValIleValAlaAlaLysAspG	663
Db	6653	AGGCATCAAGAAACAGACACACAAATCCAGATGCTCTTACTGATGTGATCTCAATGAAA	6712
Qy	663	IySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA	681
Db	6713	CCTACGATTAACGGGATCTCACTTGGAGCAACAAACACATCAAGACCAACCAACCGTGT	6772

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5286 BP; 1786 A; 1178 C; 1073 G; 1249 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,09e-06 Length: 5286
Score: 223.50 Matches: 210
Percent Similarity: 34.23% Conservative: 173
Best Local Similarity: 18.77% Mismatches: 412
Query Match: 4.04% Indels: 325
DB: 5 Gaps: 52

US-09-721-114-2 (1-1057) x AAS73156 (1-5286)

QY 83 HisAspGln-----LysLysCysAspGluHisLysAlaSerSerProPheSer 99
DB 2236 CATGATGAAGTGTGATGTCRAATGCCATCCATCTCTCTTTAGGATGAAGAGGA 2295
QY 100 ValAlaLysPheArgTrpAspCysSerLysCysLeuAspLysLysThrSerAsp 119
DB 2296 AAGGAAATAAGGCATCAT-----ATATCCTGTATTGAAAGTTAAGCAAAACAGAA 2349
QY 120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
DB 2350 AGTATATCATCTACCC-----ACCAAGTATCATCAGGAGC 2382
QY 140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
DB 2383 CTC-----ATTGAAGCAATCAACAAAGCAATCCAAAGTTCTGAA 2421
QY 159 -----ProSerThrGlnSerSer-----GlnGlyLys 167
DB 2422 CTTCACACAATTATTTGACTCTGCCAAGAAATCAAGCAGTTTCTCATACATGCGCAGG 2481
QY 168 AsnAlaaspArgSerThrLeuProLysSerValGlnGluGly----- 181
DB 2482 CAGTCGGAAGTAAATATATGGCTGCTTCATTGAGGATGGGCCACCTCCCTTCCAAATC 2541
QY 182 -----AsnAspSerLysCysAsnAla 188
DB 2542 AAAAAATATGGAAGATGCAATGGGGAATATATGTTAAACAAATTTAGTCCAGTCT 2601
QY 189 ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
DB 2602 CCTGAGTCAGCGAATGAATGTTCCAAAGTCTCTTCAGACTCAGCC-----CTGGAA 2652
QY 209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
DB 2653 GCACCTGAAGCCACAGAGAGATGCAAAATGTAAGGAGTGTGATCTCTCGTTAGA 2712
QY 229 ValGlyAlaLeuPro-----GluValProGlnIleThrTrpHis 241
DB 2713 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATCTCATGTCCCTCAGGGAGCCACAT 2772
QY 242 IleGlu-----ValAsnGlyAlaAspGlnProPro 251
DB 2773 GCCTCAACTCGAAGAGAGAGAGAAAGCCATTGACCTCAGCGATGATCTTCGAG 2832
QY 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
DB 2833 CTAACACCAAGGGCTTGGGAGAGAAATCATTAGCCCTGTGGAAAGTGAATCATCTCT----- 2886
QY 272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys----- 284
DB 2887 -----GTTAGAGATTGTTTAAACCAAAAGACAAACCAAAAGGAAAC 2931
QY 285 -----AspProAsnProMetSerGlyLysGluArgAspGlnValAlaGlu 299

DB 2932 TTCCAAGATATACACTGAGAAAGAGACTTGTCAGGAAGAAGACGTGGGAAACATTACAT 2991
QY 300 GlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCys 319
DB 2992 AAAGTAAAG---ACGACTAGTAGCTTTCTGTTCTTCTGGTGAATGAATAATGTAATGT 3048
QY 320 AsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys----- 335
DB 3049 CTGGAGTGGTCTCAATATATATACACTCTACCGAGAAACCCAGCAAAAATTCGTATAC 3108
QY 335 ----- 335
DB 3109 CTCCTTCAACAGTATACACAAATACTAATTTACTTATAGAAATCACCTCAAGTGGAGACT 3168
QY 336 -----ArgLysThrAspLysLysLeuMet-----LysLysG 346
DB 3169 GAAACATTTCTAAACGCTTTGAAAGAAACAAACAGATATTATCTACACGACGAGCTCA 3228
QY 346 nGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgAr 366
DB 3229 GGAACACCTTCATGTGAAATCTAAAGATGT-----CCGTCAACTCTGATCAGAG 3279
QY 366 gLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSe 386
DB 3280 CTCACCACTGAAATATGACTCGCTTCGATATATCA-AAATAGGGGGCCCTAGCTCCTAC 3338
QY 386 rArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspArgSerTh 406
DB 3339 ATTACAGGAATGGCTTCGTTGAGCCAGCT----- 3369
QY 406 rIleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGl 426
DB 3370 -GTTTCTCTCTGGAAGCAATCTAAAGCTAGAGAGATTTTTCAGATAATTTAGCTAA 3428
QY 426 uAspGlyLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAs 446
DB 3429 AACACCTCTAGTGTATTGAAACAAAGAGAGAGAA- 3465
QY 446 pGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisH 466
DB 3466 -----GGCAAAAAGTTGCAAAAGTGAACACCCCTGCATAC 3497
QY 466 sThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerTh 486
DB 3498 ATCATGTGATGCTTCAGAGAAAATGTTATCCGAGAGAAAA---TCTGAAAATTTGCAACA 3554
QY 486 rGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAs 506
DB 3555 ATCCATTAAATTCAGTACAGTGGTCCCTCTAGTCTTCCAGCT---CTTTCAGAGTTAA 3611
QY 506 pValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThr-- 525
DB 3612 TATTCGAAATTCCTCAAACTAGAGAGAGTCTTGGGAGTGTACAGGAGTGGTAGGCCAT 3671
QY 526 -----AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTy 541
DB 3672 TCATTATTCTGAGTGGCAAGTGTCCCGAGAGAGATCCACATCCACAGCTGTAGGTGA 3731
QY 541 rGly-----GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAs 559
DB 3732 TGGCTCCAGTGGATCAGCCCTAGGAGAGGAGAGGG-----GACATTGGAAC 3779
QY 559 pGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys----- 575
DB 3780 CAACTGCCCAAAAATGACT---AATAAACAATTTCTCTACTCAGAGAGTCAAGTCTTTC 3836
QY 576 -ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu----- 591
DB 3837 CCTTACTCCAGCATTCATAAACTACAGCTTGGTGAGGAGACTCAGTCAAGTGAACCAAAA 3896
QY 592 -----GlnSerLeuProLysLys-----LysLysLysGlnLysLeuGluVa 605
DB 3897 CTTAGAGAGTCTGCACTCTGAAACAGAGATTTACCTCAAGAGAGTCAAGGAGCAATAT 3956

Qy 605 lThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAl 625
Dy 3957 GACAGAGACGAGAG-----GC 3974
Qy 625 aLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI1 645
Dy 3975 TGAAGATGAATGACAGAGTCAGCTTGGGATCAACCTTCCTCTGTAAGGAACACAA-- 4032
Qy 645 eGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspGlySerAs 665
Dy 4033 ----AATAAACCAACTGGATGAC-----CTAGTAACGGGGAATAAGATCTTC 4079
Qy 665 pTyr-----AlaSerSerValPheAspThrAsn----- 674
Dy 4080 AGTTAAACACAGATGGCAGCCATGCTTAAGCCAGCAGCAAAATTCCTCCAGTAAGATGT 4139
Qy 675 -SerGlnGlnLysSerLeuAlaSer-----GlnSerThrGlnLysGluLeuGlnI 691
Dy 4140 AAGCCCCAGAGACATGACTACTATCTTCCCCCAAGGGAAGCAGATCTGGCTTTGA 4199
Qy 691 yHisLeuAlaLeuThrThrGlnGlu-----SerProHisProGlnAs 705
Dy 4200 CCATTATCTCTGGCACAGTGGAGTCAACCCACTGTTCCCTGAGCTACTCCAAATC 4259
Qy 705 nPheGlnSerThrGlnGluGln-----ThrHisLeuArgMetGluI 720
Dy 4260 TGCAGAGTCCATTGGCGAAGACAGGTGTAGTGAAGACGGAAGCATGTGAAGAAATCCGA 4319
Qy 720 u-----MetValThrIleAl 725
Dy 4320 GAACCTCTCCCATCTACTACTACTACCAACAGAGAACCTTCTACACAGCTCAGCAACCA 4379
Qy 725 aAlaSerSerProLeuPheSerHisHisAspAspGlnTyr-----IleAlaGluAlaPr 743
Dy 4380 GAAGTCTTAACAGCATTTCAACAGCAGCATCAGATGATTTAAACAGCTCTCAGAAATCACC 4439
Qy 743 oThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaTh 763
Dy 4440 ATCAACGCTGAGAAATCTTAAGAC-----GTCAC 4469
Qy 763 fThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaVa 783
Dy 4470 AGCAGCTCAGAAATTTAGTAAGAGAAATCAGGAGCCCCA-----TCACCCAT 4514
Qy 783 lAspLeuThrSerThrHisValMetGlySerSerAsnTyrAlaSerArgGlnProVa 803
Dy 4515 CACATTCACAGCTCAGGAGCAGCAATCTCTGACAACTCAGAGAGGCTGAGCCCTCC 4574
Qy 803 lIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPh 823
Dy 4575 TTTT---CCACTGGAGCT---GCACAGAA----- 4599
Qy 823 eProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgAsnAlaGlyI 843
Dy 4600 ----TCTAGATAGCAGCTCAGCTGGCAGTTTCTGACGAAACAAAGAGTGTCTTCATC 4655
Qy 843 nValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspPr 863
Dy 4656 TCTG-----GAGTGGGAACCTGAGCCACACCTCTATCGT-----TC 4691
Qy 863 oSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGl 883
Dy 4692 AAAGAGTTTAAAGACATTAATGTTTCATGGCGATCTACTACGAAACAGC----- 4740
Qy 883 nLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSertyrGl 903
Dy 4741 ----CATCTCCAAAGTCAGGAGGCCATTTTCTGAAAGCATCTCTATT----- 4788
Qy 903 ySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHi 923
Dy 4789 -GACAATGCCCTGAGTCGACTGACCTTGGGAATGAATTTCTCTCAACAATGGGTACAG 4847

Qy 923 sAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGl 943
Dy 4848 TCGAAGATTCAGATCTTTTCTGAACCTCCCTCTGTGAT-----GGAAAT-- 4893
Qy 943 nLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyVa 963
Dy 4894 -----GAAAGTTGGCTTATCGCAGCGGACAAAACAGGTCCAGGCTGTGCAAT 4943
Qy 963 lSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSe 983
Dy 4944 ATCTATA-----TACAGACCTTATCGACTA 4967
Qy 993 rGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerAlaAs 1003
Dy 4968 TGGGATCTTTGGGAAGAACCAACAGTTAGCTTCTTGAATAATGTAAAGAGGTCA-- 5022
Qy 1003 pPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetVa 1023
Dy 5023 ----CTTACACAGGAAGATTATGCAACCAAGTTTCTTAAGAACCCTGGC----- 5070
Qy 1023 lHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
Dy 5071 -----TTCTTAAAGATGATTG-----AGAACCTT 5097
RESULT 11
ABL68623
ID ABL68623 standard; DNA; 10452 BP.
XX ABL68623;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:6960.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 22-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.

```

PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX
XX WPI; 2002-189264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 6960; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
XX tumour
XX
XX SQ Sequence 10452 BP; 3702 A; 1658 C; 2212 G; 2880 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5.61e-06 Length: 10452
XX Score: 223.00 Matches: 172
XX Percent Similarity: 34.31% Conservative: 108
XX Best Local Similarity: 21.08% Mismatches: 312
XX Query Match: 4.04% Indels: 225
XX DB: 6 Gaps: 33
XX
XX US-09-721-114-2 (1-1057) x ABL68623 (1-10452)
XX
XX Qy 62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
XX Db 1547 AAGCTCATCTTGGATTGGAGAGAGACTTAATTCGAGTTTCGAGCGATGGATGCTGTA 1606
XX
XX Qy 82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
XX Db 1507 ---AACCAAGAGAAAATACCAAGAGACATAAGTCATAGATGCTAGATTGTAACCAAAA 1663
XX
XX Qy 101 -----AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
XX Db 1864 GCACGAAAAGGAGAAAACCTTGTCTTGGAAAAGAGGATATTTCAAAGTCGAGACT 1723
XX
XX Qy 114 LysLeu----- 115

```

```

Db 1724 AAACITTTCAAGAAAACAGGTAGTAGTGTAGCAGCATGTCATCAGAAATGTTCCACAGAGGAA 1783
Qy 116 -----LysThrSerAspAsnGlyThrAlaPro 124
Db 1784 CAAAGAACAAATAAAGTAGTACCAGTGTGTGAACATAGAAATCTGATAGAAAAGAAACCT 1843
Qy 125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
Db 1844 CAATATGAACCTGCC-----AACCTTCTGAAGATTAGACATGATATGTTGTCT 1894
Qy 145 SerThrPheValProAlaSerValGlySerGln-----LysValSerPro 159
Db 1895 -----GTTCTTCTCAGTTCAGAGACATTTTGGAGAAATCTTGAGACTGCTATG 1945
Qy 160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db 1946 GAAGTTTCAGAGTTTCAGTTGATCATCAAGGGATGGCAGCAGTGGAACTGAACAAGAGTG 2005
Qy 180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn-----GlyAlaAlaGlu 197
Db 2006 GAGAGTTTCATCTGTAAATTAATATATTTCTTCAABAGACACAGAGGAGGTATTAATCA 2065
Qy 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db 2066 AAAAAGTACAGCTAAAGTAAACAAAGAAATTA-----TATGTTAAACTC 2107
Qy 218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db 2108 ACTCTGTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2167
Qy 238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257
Db 2168 -----GATAAAGATGCTATATAAAGTTGTGTCTGAACCCCAAGTTAGAG 2212
Qy 258 GluValValLeuLysArgAsnGluAspGluAsnGly-----LysThrGluGluThr 274
Db 2213 AAATGTGGACTGGACAGAAACAGTGAATAGACATTTGGTTGAAATGAAGTTTCA 2272
Qy 275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db 2273 TTACTTTTAGAGGAATCTGATCTTCGAGATCCCCAGCTGTAAGACTACACCCCTGAGG 2332
Qy 295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProLysSerGlyGlnLys 314
Db 2333 CGACCGACA-----GAAACTAACCCCTGTAACATCTAATCTCA 2368
Qy 315 CysGluGlnIleCysAsnGluProCysGluValValLeuLysArgSerLysSer 334
Db 2369 GATGAAGAA---TGTAATGAACAGTAAAGAG-----AAACAAAACATATCAGTTCCA 2419
Qy 335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db 2420 GTGAGAAAAAAGAGATAAGGCT----- 2440
Qy 355 GlnAlaAspValSerAspAlaLysLeuCysArgLysProLysLysVal-----Arg 372
Db 2441 -----AATTTCTTCAGAGTGTCTATGATAATCTTAAAGCTTAATAATTCGCAAACTCT 2494
Qy 373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg-----SerAsp 389
Db 2495 AAGCAATCAGAGACTGTGATCAAAATTCAGATTCTGTGAATATGTCAGAAATCTCCCAAA 2554
Qy 390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409
Db 2555 GAGGTGAGCAGGATGAGTCAAGTCTTCTTCAGATACATGATATTAATTAATTCATACA 2614
Qy 410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db 2615 AACCATAGACTTTGTATGATTTAAAGACTCAG-----GCGGGAAGAGAT----- 2659
Qy 430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp----- 446

```


PA (ITOL/) ITO L Y.
 PA (SHER/) SHERMAN B K.
 XX Lalgudi RV, Ito iY, Sherman BK;
 XX WPI; 2002-163647/21.
 XX Novel purified corn tassell-derived polynucleotide useful for determining
 PT altered gene expression, to recover regulatory elements and to follow
 PT inheritance of desirable characteristics through hybrid breeding
 PT programs.
 XX Claim 1; SEQ ID NO 2759; 201pp; English.
 XX The present sequence describes a purified corn tassell-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70427 to ABL76833. The cdps sequences
 CC encode corn tassell-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful in
 CC the evaluation, and alteration of desired characteristics associated with
 CC growth and development, disease resistance, environmental adaptability,
 CC quality and yield, and as molecular markers for studying inheritance of
 CC multigene traits in a plant breeding program. (I) can be used to produce
 CC a tassell-specific profile of gene transcription, a transcript image, to
 CC clone regulatory elements for use in transformation vectors, to express a
 CC polypeptide, to identify, isolate or extend identical or related corn
 CC tassell nucleic acid sequences from DNA libraries, in nucleic acid
 CC hybridisation or amplification technologies, as query sequences to
 CC determine homology of known sequences, as probe for use in Southern or
 CC Northern hybridisation, and to identify the presence of and/or to
 CC sequences
 XX Sequence 279 BP; 70 A; 74 C; 73 G; 56 T; 0 U; 6 Other;
 SQ
 Alignment Scores:
 Pred. No.: 5.09e-08 Length: 279
 Score: 222.50 Matches: 55
 Percent Similarity: 64.58% Conservative: 7
 Best Local Similarity: 57.29% Mismatches: 25
 Query Match: 4.03% Indels: 10
 DB: 6 Gaps: 4
 US-09-721-114-2 (1-1057) x ABL73395 (1-279)
 QY 871 AsnTyrGlyThrSerArgAsnGlnMetGlu-----SerGlnLeuHisAsnSergin 888
 DB 2 AACTACCAAGAGCTAATAAGCGGCANATGGAGCTTCMAAGCAAGCTCTCGGCTCGCAG 61
 QY 889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
 DB 62 TACACACAGCATGACCATTAACATGCGTCCAGCAGCAGTATATGAAGCCAA----- 115
 QY 908 GlyLysileProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
 DB 116 -----CCGCTGANACTGGAGAGACTTCTCGCGTTCATTCAGCAAGACTTGCGC 166
 QY 927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuGlySerLeuGlyLue 946
 DB 167 AGGCCTTTACGNCNCTACCCCTCGANTCGTGTGCTCGGTTTCATTCGTCGACGAGAGATC 226
 QY 947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLeuGly 962
 DB 227 GCAAACTGGTCTGG-GNATCGGGCCCGCAGTCTGGGTANAGACTAGT 273
 RESULT 15
 AAS80591
 ID AAS80591 standard; cDNA; 5954 BP.
 XX
 AC AAS80591;
 XX

DT 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #16395.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI: 2001-639362/73.
 XX P-PSDB; ABG16404.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 16395; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5954 BP; 2007 A; 1302 C; 1219 G; 1426 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.76e-06 Length: 5954
 Score: 221.00 Matches: 213
 Percent Similarity: 33.71% Conservative: 172
 Best Local Similarity: 18.65% Mismatches: 410
 Query Match: 4.00% Indels: 248
 DB: 5 Gaps: 53
 US-09-721-114-2 (1-1057) x AAS80591 (1-5954)
 QY 83 HisAspGln-----LysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
 DB 2841 CATGATGAAGTGTTGATGTCTCAATGCCATTCACACTCTCTTTAGGATCAAGAGGA 2900
 QY 100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuThrSerAsp 119
 DB 2901 AAAGGAAAATAAGGCATCAT-----ATATCTCTATTGAAAGCTTAAGCAAAACAGAA 2954

Qy 120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
:: :: :: ||| |||||
Db 2955 AGTATATCACTAGCC-----ACCACTGATCACGAGC 2987

Qy 140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
:: :: :: ||| :: :: |||||
Db 2988 CTC-----ATTGAAGCAATCAAAAGCAATTCCAAAGTTTCTGAA 3026

Qy 159 -----ProSerThrGlnSerSer-----GlnGlyLys 167
:: :: :: ||| :: :: |||||
Db 3027 CTTGACACAATTTATTCTACTTGCCCAAGAAAATCAAGCAGTTTTCTCATACATGGCAGG 3086

Qy 168 AsnAlaaspArgSerThrLeuProLysSerValGlnGlucly- - - - - 181
:: :: :: ||| :: :: |||||
Db 3087 CAGTCAGAAAGTAATAATAGGTGCTTTCATTGAGGAATGGGCACCCTCCCTTCCAATC 3146

Qy 182 -----AsnAspSerLysCysAsnAla 188
:: :: :: ||| :: :: |||||
Db 3147 AAAAAAATGTGGAAGATGCGAATGGGNACTATATGTTAAACAAATTAGTCCCAGTTCT 3206

Qy 189 ProSerGlyLysAsnGlyAlaLaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
:: :: :: ||| :: :: |||||
Db 3207 CCTGAGTCAGCGAATGAATGTTCCAAAGTCTCTTCAGACTCAGCC-----CTGGAA 3257

Qy 209 GlyProAlaGlnAsnTyraSpValAlaLaalaasnValSerGluAspAsnThrSerValaSp 228
:: :: :: ||| :: :: |||||
Db 3258 GCACCTGAAGCCACAGAGAGATGACAAATGTAAAGACGAGTCACTCTTCGTTAGA 3317

Qy 229 ValGlyAlaLeuPro-----GluValProGlnIleThrTrpHis 241
:: :: :: ||| :: :: |||||
Db 3318 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATGTCACTGCCCTCAGGGAGCCACAT 3377

Qy 242 Ileglu-----ValAsnGlyAlaAspGlnPro 251
:: :: :: ||| :: :: |||||
Db 3378 GCCTCAACTGGAAGAGAGAGAAAAGCCATTGACCTCAGCATGGAATGCTCTGAG 3437

Qy 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGlnLysThr 271
:: :: :: ||| :: :: |||||
Db 3438 CTAAACCAAGGGCTTGGGAGAGAATCATTAAGCCCTGTGGAAGTGAATCATCT----- 3491

Qy 272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys----- 284
:: :: :: ||| :: :: |||||
Db 3492 -----GTTAGAGATTGTTCTTTAACCAAAAGACACACCAAAAGGAAAAAC 3536

Qy 284 ----- - - - - - 284
:: :: :: ||| :: :: |||||
Db 3537 TTCCAAGAATACACTGGAAGAGCGTAAATGCGCTGCCCTCAGGAGAGTGATTGCT 3596

Qy 285 -----AspProAsnPro-----MetSerGlyLysGluArgAspGln 296
:: :: :: ||| :: :: |||||
Db 3597 CTTTCAAATGAAGACCCCTTTACCTTTTGTCTCAGACTGTGAGAAAAGAACCTGGGAAA 3656

Qy 297 ValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGlu 316
:: :: :: ||| :: :: |||||
Db 3657 ACATTACATAAAGTTAG--ACGACTAGTACGTTTCTGTTCTGTTGATGAAGATAAT 3713

Qy 317 GlnIleCysAsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys-- 335
:: :: :: ||| :: :: |||||
Db 3714 GTAAAAATGCTGGAGGTGGTCTCAATATATTACACTCTACCGAGAAACCCAGCAAAAAA 3773

Qy 335 ----- - - - - - 335
:: :: :: ||| :: :: |||||
Db 3774 TTTGTAACTCTTCAACAGTATACAAAAATACTAATTACTTAGAATCACCTCAA 3833

Qy 336 -----ArgLysThrAspLysLysLeuMet----- 343
:: :: :: ||| :: :: |||||
Db 3834 GTGAGACTGAACAFITCTCTAAGCTTTTAGAAAAAGACAAACAGATTTATTCTACAGA 3893

Qy 344 -LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaaspValSerAspAlaLysLe 363
:: :: :: ||| :: :: |||||
Db 3894 GAGCAGTCAGGAACACCTTCTATGTGAATAATCTAAAGATGT-----CCGTCAACTCT 3944

Qy	363	uCySArgArgLysProlylsValArgLeuLeuSerGluIleAlaAsnAlaAsnGlnVa	383
Db	3945	GATCAGACGGCTCCACTGAATAATATGCTCGCTTCGATATATCA--AATAGGGGGCCCCCT	4003
Qy	383	lGlUASerArgSerAspGluValHisArgGluAsnAlaAspProCysGluAspAs	403
Db	4004	AGTCTCTACATTACAGGAATGCCTCTGTGAGGCAGCT-----	4043
Qy	403	pArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisTh	423
Db	4044	---::: ---GTTTCTCTCTCGAAGAGGAATCTAAAGCTTAGAGAGATTTTTTCAGATAA	4093
Qy	423	rValSpGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspVa	443
Db	4094	TTTAGCTTAACAACCTCTAGTGATTCAGAAAACAAGAGAAGAA-----	4139
Qy	443	lValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySe	463
Db	4140	---::: ---GGCAAAAGTTGCAAAAGTGA AAC	4162
Qy	463	rValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProTh	483
Db	4163	CCTGCATCTTCATTGATGCTTCAGAGAAAAATGATCCGAAGAAAA---TCTGMAAA	4219
Qy	483	xAlaSerThrGlnHiEaspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHi	503
Db	4220	TTTGCAACAATCCATTAAATCAAGTAACAGTGTCTCTCTAGCTCTCCAGCT---CTTTC	4276
Qy	503	sLysThrAspValCyeGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGl	523
Db	4277	AGAAATTAATATTGGAAATTTCCCAAATTCAGAACAAAGTTCTTGGAGTGTCACGGGAGTGG	4336
Qy	523	vLysThr-----AlaGlyLeuSerLysGlyLysThrHisSerAlaAlase	538
Db	4337	TAGAGCCATTCCATTACTGGAAGTGGCAGTGTCCCAGAAAGATCACACATCCACAGC	4396
Qy	538	rThrLysTyGly-----GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSe	556
Db	4397	TGTAGTGATGGCTGCCAGTGGATCACAGCCTAGGAGGCAGAGG-----GA	4444
Qy	556	xAlaGluAspGlnCysGlnMetGlnThrGluAsnSerValLeuSerHisSerAlaLys--	575
Db	4445	CATTGGAAACCACCTGCCAAAAATGACT---AATAAACACTTTCTCATCTCAGAGAGTCA	4501
Qy	576	-----ValSerProAlaGluHisApIleGlnIleMetSerAspLeuHisGlu--	591
Db	4502	AGTCTTTGCCCTTACTCCAGCATTTGCATAAACTACAGCTTGGTGAGAGACTCAGTCAGA	4561
Qy	592	-----GlnSerLeuProLysLys-----LysLysLysGlnLy	602
Db	4562	TGAACCAAACTTAGAGAGTCTGACAGTCTGAACCAAGAGAAATTACTCAAGAGAGTCAGGA	4621
Qy	602	sLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGl	622
Db	4622	GGCAAAATATGACAGAGACAGGAAG-----	4646
Qy	622	uLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIl	642
Db	4647	-----GCTCAAGATCAAAATGCAAGAGCTCAGCTTGGGATCAACCTTCACTTCTCGAAGG	4699
Qy	642	eAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAs	662
Db	4700	AACAAA-----AATAAAACCAACTTGGATGAC-----CTAGTAAGGGGGGAAAA	4744
Qy	662	pGlySerAspTyr-----AlaSerSerValPheAspTh	673
Db	4745	TAGATCTTCAGTTAAACACAGATGGCGAGCCATGTCCTAAAGCCACAGAGAAAATTC	4804
Qy	673	xAsn-----SerGlnGlnLysSerLeuAlaSer-----GloSerThrGlnLysGl	688
Db	4805	TAAAGATGTAAAGCCCCCAAGAGCATGTAGTACTATCTTCTCCCCCAAGTGGAGACGATC	4864
Qy	688	uLeuGlnGlyHisLeuAlaLeuThrThrGlnGlu-----SerProHl	702

Db 4865 TGGCTTTGACCAATTATCTCTTGGACAGTGGAGTGCACACCCACTTTCCTCGAGCCTAC 4924
QY 702 sProGlnAsnGlnSer-----ThrGlnGlnGlnThrHisLeuAr 717
Db 4925 TCCAAATCTGCAGAGTCCATTGGCGAAAGCAGTTGAGTGAGAACAGAAAGCATGTGAA 4984
QY 717 gMetGluGlu-----MetVa 722
Db 4985 GAATCCGAGAACCTTCTCCCAATTACTGTACTACCCACAGAGAACCTTCTACACAGT 5044
QY 722 lThrIleAlaAlaSerProLeuPheSerHisHisAspGlnTyr-----IleAl 740
Db 5045 CAGCAACCCAGAGTCTTACAGCATTTCACACGACATCAGATGAGTTTAAAAACGTCTC 5104
QY 740 aGluAlaProThrGlnHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGlnPh 760
Db 5105 AGAATCACCATCAAGCATGAGAAATCTTAAGAC----- 5138
QY 760 eLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyI 780
Db 5139 ----GTCACAGCAGCTCAGAAATTAGTAAGAGAATCAGGAGCCCA----- 5180
QY 780 eGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerAr 800
Db 5181 -TCACCCATCACAATTCACCAGCCTCAGGGAAGCAGAAATCTCTGCAATCAGAGAGGCT 5239
QY 800 gGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAl 820
Db 5240 GAGCCCTCTCTTT---CCACTGGAGCT---GCACAGAAA----- 5273
QY 820 aArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAs 840
Db 5274 -----TCTAGAGTAAGCAGTCCACTGGCCAGTTTCTCAGCAACAAAGGAG 5320
QY 840 nAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMe 860
Db 5321 TGCTTCATCTCG-----GAGTGGAACTGAGCCACACCTCTATCGT--- 5363
QY 860 tMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMe 880
Db 5364 -----TCAAGAGTTTAAAGAGCATTAATGTTTCATGGCGATCTACTACGAAAAAGC-- 5414
QY 880 tGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerTh 900
Db 5415 -----CATCTCCAAAGTCAGGAGCGCCATTTTCTGAAAGCACTTCTAT 5461
QY 900 rSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHi 920
Db 5462 T-----GACAAATGCCCTGAGTCGACTGACCTTGGGAATGAATCTCTGTCAACAA 5512
QY 920 sGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySe 940
Db 5513 TGGGTACAGTCGAAGATTCAGATCTTTTCTGAACCTCCCTCCTGTGAT-----GAAA 5566
QY 940 rLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLy 960
Db 5567 T-----GAAAGTTGGGCTTATCGCAGCGGACAAACAGGTCCCGAG 5608
QY 960 sLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAl 980
Db 5609 GTCTGCAATATCTATA-----TACAGACC 5632
QY 980 aLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSe 1000
Db 5633 TATCGACTATGGGATCTTTGGGAAGAACACACAGTAGTCTTTCTTACAAATCTAAAGAG 5692
QY 1000 rSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGl 1020
Db 5693 GTCA-----CTTACAAAGAGAGATTAATGGAACCAAGTTTCTTAAGAACCCCTGG 5743
QY 1020 yLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAs 1040

Db 5744 C-----TTCTTAAAGATGATTG-----AGGAA 5767
QY 1040 nPro 1041
Db 5768 CCT 5771
Search completed: May 17, 2004, 05:34:03
Job time : 1320 secs

1	298	5.4	8548	4	US-09-415-945-2	Sequence 2, Appl
2	298	5.4	17341	4	US-09-415-946-1	Sequence 1, Appl
3	214	3.9	4584	4	US-09-845-917A-51	Sequence 51, Appl
4	214	3.9	4749	4	US-09-845-917A-55	Sequence 55, Appl
5	214	3.9	5009	4	US-09-845-917A-52	Sequence 52, Appl
6	214	3.9	5010	4	US-09-845-917A-56	Sequence 56, Appl
7	214	3.9	5072	4	US-09-845-917A-2	Sequence 2, Appl
8	214	3.9	5073	4	US-09-845-917A-1	Sequence 1, Appl
9	214	3.9	9842	4	US-09-845-917A-30	Sequence 30, Appl
10	214	3.9	10443	4	US-09-845-917A-35	Sequence 35, Appl
11	214	3.9	13414	4	US-09-845-917A-37	Sequence 37, Appl
12	211	3.8	7474	4	US-09-845-917A-36	Sequence 26, Appl

```

OTHER INFORMATION: exon 2
NAME/KEY: intron
LOCATION: (4336)..(4447)
NAME/KEY: exon
LOCATION: (4448)..(4623)
OTHER INFORMATION: exon 3
NAME/KEY: intron
LOCATION: (4624)..(4703)
NAME/KEY: exon
LOCATION: (4704)..(4823)
OTHER INFORMATION: exon 4
NAME/KEY: intron
LOCATION: (4824)..(4902)
NAME/KEY: exon
LOCATION: (4903)..(4956)
OTHER INFORMATION: exon 5
NAME/KEY: intron
LOCATION: (4957)..(5045)
NAME/KEY: exon
LOCATION: (5046)..(5307)
OTHER INFORMATION: exon 6
NAME/KEY: intron
LOCATION: (6308)..(6447)
NAME/KEY: exon
LOCATION: (6448)..(8065)
OTHER INFORMATION: exon 7
NAME/KEY: intron
LOCATION: (8066)..(8300)
NAME/KEY: exon
LOCATION: (8301)..(8648)
OTHER INFORMATION: exon 8
US-09-415-946-2

Alignment Scores:
Pred. No.: 1,03e-17
Score: 298.00
Percent Similarity: 35.53%
Best Local Similarity: 21.21%
Query Match: 5.39%
DB: 4

US-09-721-114-2 (1-1057) x US-09-415-946-2 (1-8648)
Qy 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db : : : : :
5090 TCTAATGTTAGGACGACGACGAGAGCTTGTGAGTCGGGAGCAG-----GTAGCA 5140
Qy 218 AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu 234
Db : : : : :
5141 AACAACGAGATCTAAGAGAAAAAATAAACAACCACTCGATGATATTAGCAGCTCGAAAGAG 5200
Qy 235 ValProGlnIleThrTpHisIleGluValAsnGlyAla----- 247
Db : : : : :
5201 AAACAAATGTCGATCAGGCTGTGACAACTTCGGCTCATCTGAAATTCGTGGTGGTT 5260
Qy 248 ----AspGlnProProSerThrProLysLeuSerGluValValLysValLysArgAsnGluAsp 266
Db : : : : :
5261 GAGGATACACACCACTAAGGCAACCAAGAAATCATAAAGGCAATTCGCGGCTCTGATCGAATGC 5320
Qy 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db : : : : :
5321 GATAACGGGTCATCAGAAAGTATAAATCTGCT----- 5353
Qy 287 AsnProMetSerGlyLysGluArg-----AspGlnValAlaGluGlnCysAsnLeu 303
Db : : : : :
5354 -----ATGAGTGGTGTGACGCTAGGAAATCTCGAAGGTTCTCTACTCTCAGTGAGTTG 5407
Qy 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db : : : : :
5408 CTTGGTAAATCAAAAACCACTGGTGTGTAGTAACATCAAGAAA----- 5449
Qy 324 GluGluValValLysLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db : : : : :

```

```

5450 GAAGAGTCTGCTTTGCAAGAGGAATCAGTTAGAGTCGAAAA-----AGAANGTTGTTA 5503
Qy 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp----- 360
Db : : : : :
5504 CCGTGAACAACATTTATCTCAGCCGGATATTGACTACATGGGGTCAACCTCTCTCAAAATGCT 5563
Qy 361 AlaLysLeuCys----- 364
Db : : : : :
5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATGAGGGTTTGACAGA 5623
Qy 365 -----ArgArgLysProLysLysValArgLeuLeuSerGluLeuIleAsn 379
Db : : : : :
5624 ACTCCATTTAAGGTTAAGCAGAGAAACAGAGATTTCAAGTTGTTGACGAGTTTGTACCA 5683
Qy 380 AlaAsnGlnValGluAspSerArgSerArgSerGluValHisArgGluAsnAlaAlaAspPro 399
Db : : : : :
5684 TCACCTTCCTGTGAAAACCTTCAAA-----GAAGGTATCAAGGACATGATGATGATCTCT 5737
Qy 400 CysGluAspAspArgSerThr-----11ePro 408
Db : : : : :
5738 AGTAAG-----AGATCAACTCTCGGCACCTTTATTTCACCTGGAACGATTCCTGCTCT 5791
Qy 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db : : : : :
5792 TGTCTCT-----CCGGGTACTCAGAGAAC-----GAGAGGAG 5824
Qy 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db : : : : :
5825 CTCAGTTTACCCAGAGAGACAAAGAG-----CCTGTAATCGAATATGGGAG 5875
Qy 449 SerLeuMet-----AsnTrpLeuAsnGlyLysLys-----LysArgThrGly 462
Db : : : : :
5876 AGCAGTGTGATCAGTTTGTAGTAACGGCATCTGATGGAAGTCAAGTTAACTCGCATACTGCT 5935
Qy 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysValThrPro 482
Db : : : : :
5936 CCTTCATCAACACAGATATCCAAACTCGAGACTTATTGAATGGGAAGGGTGGCGGT 5995
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLysLeuAsp----- 499
Db : : : : :
5996 TTATTGTCAACCGTTTGTGCTTCAGATGATATTTTCAGAAAAATATCTCTCAGGTAAAT 6055
Qy 500 -----ThrAsnMetHisLysThrAspValCysGlnHisValSerGluLeuSer 515
Db : : : : :
6056 GATAAGCCGATACATCTTTCGATTTGCAAGCAATGATATTGTGAGGTCAAGACACGG 6115
Qy 516 ThrGlnArgCys-----SerSerLysGlyLysThrAlaGly---Leu 528
Db : : : : :
6116 GAACCAAACTGCTTCGAGATTTTAGTTCCTCTCTTAATCCAGCTCAGGTGGATGTTG 6175
Qy 529 SerLysGly-----LysThr 533
Db : : : : :
6176 AGAAGTGGAGTAGATATTGTTGACTTCAGAAAACCAACCAATACAAACAGATCGTCT 6235
Qy 534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGluAsnIleHis 553
Db : : : : :
6236 TTCTCGAAGTGAAGCTAAGATACCCCTCTCTCTACTGAAAGTTGCGGATTTATCTCGG 6295
Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db : : : : :
6296 GTGCTGCAAAAGGTACAAGCTTTGTCATCTATTTTCTTCAATCGCCATGCTTCTCATTC 6355
Qy 573 I----- 573
Db : : : : :
6356 TGTTCACACTGTCACCTATTTCTGTATATTCCAAACCTGGTGTCTCAGTGATGGAT 6415
Qy 574 -----Ala-LysValSerProAlaGluHisAspIleG 584
Db : : : : :
6416 CAGGATGATACGGGATTTGATTCTTGGCAGGATGCTTCTGTCAGATAGAAAAGGGA 6475
Qy 584 InIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG 604
Db : : : : :
6476 AGACTGTTATGTTCCAGAGACATCATGAGACCAAGAACCAAGTCCAGATAGAAAG 6535

```



```

; NAME/KEY: intron
; LOCATION: (4336)..(4447)
; NAME/KEY: exon
; LOCATION: (4448)..(4623)
; OTHER INFORMATION: exon 3
; NAME/KEY: intron
; LOCATION: (4624)..(4703)
; NAME/KEY: exon
; LOCATION: (4704)..(4823)
; OTHER INFORMATION: exon 4
; NAME/KEY: intron
; LOCATION: (4824)..(4902)
; NAME/KEY: exon
; LOCATION: (4903)..(4956)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (4957)..(5045)
; NAME/KEY: exon
; LOCATION: (5046)..(5307)
; OTHER INFORMATION: exon 6
; NAME/KEY: intron
; LOCATION: (5308)..(5447)
; NAME/KEY: exon
; LOCATION: (5448)..(8065)
; OTHER INFORMATION: exon 7
; NAME/KEY: intron
; LOCATION: (8066)..(8300)
; NAME/KEY: exon
; LOCATION: (8301)..(8648)
; OTHER INFORMATION: exon 8
US-09-415-946-1

Alignment Scores:
Pred. No.:      3,44e-17      Length:      17341
Score:          298.00       Matches:      225
Percent Similarity: 35.53%   Conservative: 152
Best Local Similarity: 21.21% Mismatches:    360
Query Match:     5.39%     Indels:       324
DB:              4         Gaps:        50

US-09-721-114-2 (1-1057) x US-09-415-946-1 (1-17341)

Qy  198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProalagInAsnTyrAspValala 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5090 TCTAATGTTAGAGCAAGACGAAAGAAGCTGTGTAGTCGGAGCAG-----GTAGGA 5140

Qy  218 AlaAsnValSerGluAsp-----AsnThrSerValAspValGlyAlaLeuProGlu 234
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5141 AACACACAGACTTAAGAAAAAATAACAACCATCGATGGATATTAGCAGTCGAAAAGAG 5200

Qy  235 ValProGlnIleThrTpHisIleGluValasnGlyala-----                247
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5201 AAACAAAATGTGGATCATCGGTGTGCACACGTTCCGGCTCACTGAAATGTCTGTGTGTT 5260

Qy  248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluasp 266
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5261 GAGGNTACACCACCTTAGGCCAACCRAGAATCATAAGGCATTCGCGTCTGATGGAATGC 5320

Qy  267 GluasnGlyLysThrGluGluThrLeuValalaGluGlnCysAsnLeuThrLysAspPro 286
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5321 GATAACGGGTGCATCAGAAAGTATAAAATCTGCT-----                    5353

Qy  287 AsnProMetSerGlyLysGluarg-----AspGlnValalaGluGlnCysAsnLeu 303
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5354 -----ATGAGTGGGTTCGACGGTAGGAAATCTCGCAAGGTTCTGTACTCAGTGAGTTG 5407

Qy  304 ThrLysAspProLysProvalserGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5408 CTTGGTAATACAAAACACCGAGTGCTGGTAGTAACATCAGAAA-----          5449

Qy  324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLeuMet 343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  5450 GAGAGTCTCTTTTGAAGAGCAATCACTTAGAGTCTGAAA-----AGAAAGTTGTTA 5503

```

QY 604 luValThrArgGluLeuGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
 DB 6536 AGACTACGACTAGAGAGCAAAACAC---GATGATATTCCTAATGAGATAGTAGGCTCA 6592
 QY 624 euAlaIysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
 DB 6593 TGGCCAAAACCCAGTACGAGAGGTCTCTCCCGCAAAAGAGAAGATGTTAGCAAAAC 6652
 QY 643 snArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspG 663
 DB 6553 AGCCATCACAAGAAACAGCACACAAATCCAGATGCTCTACTGATGATCTCAATGAA 6712
 QY 663 lysSerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-----LysSerLeuA 681
 DB 6713 CCTACGATAACGGGATCTCACTTCGAGACAAACACATCAAGACCAACCAAAACCGGTGA 6772
 QY 681 laSerGlnSerThrGlnLysGluLeuGlnGlnHisLeuAlaLeuThrTrpGlnGlnSerP 701
 DB 6773 GTAGCAACGCAAGGAGGAAGAA-----CATTTTCCTATCGGAGACAGCGCAACT 6823
 QY 701 roHis-----ProGlnAsn-----PheGlns 708
 DB 6824 CTCATGACTTCTCCCAATAAGTCAGCGCTATGTGCTTCTTCGTTTGGGATCTTTCCTC 6883
 QY 708 erThrGlnGlnGlnGlnHisLeuArgMetGluGluMetValThrIleAlaIaSerS 728
 DB 6884 CTACCCCAAGAAACCGA-----GCCAGCT 6907
 QY 728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
 DB 6908 CCATCCGGTTTCTCGTCACAACTGTCAGTGGCTTGGGAATTTGCCAAGTGTG-----G 6961
 QY 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
 DB 6962 GTAATCAGAACCTTCTCCATCTCTCAATTTCCGGTATTACGTGCT----- 7005
 QY 768 roAlaAlaThrCysGlyAla-----GlnPheArgProGlyIleGlnAlaValAspL 785
 DB 7006 --TGTGATCTTTCAGAGTGTTCCTATCATCATACAGAA----- 7041
 QY 785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
 DB 7042 --GAAGCTTCTCATCAATTTGGCCATCTTCC-----ATGATAC 7078
 QY 805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
 DB 7079 CACCACAGAGTCAATACAGCCAGCTTCTTTTAAATATTAAATCATGCTCAACAAAT---CCGG 7135
 QY 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
 DB 7136 GTACGCTTTCA-----CAGGCACTCAACATGAATATACATGGAACCTCAACTTTGTG 7189
 QY 845 alLeuTyrProLysGlu----- 850
 DB 7190 CTGCCACGGGAAGCAAAATATGGGCTTAATCCAGAAATTTTCATTTTGGCTGCACAAATG 7249
 QY 851 -----SerM 852
 DB 7250 CTGCTGGGTAGTAGTAGTAGTAGAGCCCAATAGATAACTTTTCTAGTAGAGAGCTCTA 7309
 QY 852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro----- 870
 DB 7310 TACCGGCATTGATCTACTACGCTTCTGGATCTCTGGCTCGAGGTCAACGACTCCCGCTG 7369
 QY 871 --AsnTyrGlyThrSer-----SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
 DB 7370 ACCAACCGGAACCACTAAATTTACTAAAGACATTTTCCGCCAGCCCAACCGCTCCAAAG 7429
 QY 888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
 DB 7430 AGTTTATAGAGCTTCAACAGCGGACTCTAGTAAGTCAAGCTACTCAACTAAG----- 7482
 QY 908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928

DB 7483 --CAGATACCT-----TTTGATCTATACAGCAAGAGATTACACACAGACGCTTCCCGGA 7534
 QY 928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
 DB 7535 AGAGTTTCCCATCATCTCCACCTATTGGAGC-----TCTTCATCTTCAATTTCAAAATG 7598
 QY 947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
 DB 7589 CTTTCATGGAGTCTCTCAT----- 7605
 QY 967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
 DB 7606 -----CATCAAGAGAGAAACCAAGAGAGAAAGACACC-----TTTG 7642
 QY 987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp----- 1003
 DB 7643 CTCCTGTTTACATACTCATGAAAAGCCGGTGTTCGACGACGACCAACGACCAAGCGAAGT 7702
 QY 1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
 DB 7703 TCCAGCTGCTGGGAGCATCGAATTC-----AspArgPheValArgGlnAsp-----A 7729
 QY 1022 etValHisProLeu----- 1033
 DB 7730 TGAATCTCTTTGAAATTTTCATGACGATGACGATTAAGAAAGAAACAAAGAGAAAGCAG 7789
 QY 1034 -----IleC 1035
 DB 7790 AGAGCTGCATAACAATCCCTCTCGCGGAGCTCTGGAAGAACAGTCTTGGACCCATTTGTG 7849
 QY 1035 ysIleThrLeuLysAsnProAlaAspPheThrIleSerAsnAspAsnGluTyrMet 1054
 DB 7850 GCAGCGTCAATAGAACCCCTGCTGATTTCCACCATCTCTGAACTGGGAATGTTTACATG 7908

RESULT 3

US-09-845-917A-51

; Sequence 51, Application US/09845917A

; Patent No. 6653529

; GENERAL INFORMATION:

; APPLICANT: Bogaert, Thierry

; APPLICANT: Vandekerckhove, Joel

; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH

; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR

; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR

; FILE REFERENCE: P/14-1

; CURRENT APPLICATION NUMBER: US/09/845,917A

; CURRENT FILING DATE: 2001-04-30

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 51

; LENGTH: 4584

; TYPE: DNA

; ORGANISM: Caenorhabditis elegans

US-09-845-917A-51

Alignment Scores:

Pred. No.:	5,37e-10	Length:	4584
Score:	214.00	Matches:	243
Percent Similarity:	34.68%	Conservative:	170
Best Local Similarity:	20.40%	Mismatches:	414
Query Match:	3.87%	Indels:	365
DB:	4	Gaps:	56

US-09-721-114-2 (1-1057) x US-09-845-917A-51 (1-4584)

QY 14 ValGlyThrAsnCysMet-----LeuAlaArgGlyGlyThrGlyAla 27

DB 247 CTGGGTCTCGACTGCTCGAAACTCACCAAAACCGATATCGACGCGGAAACTTGGGTGCA 306

QY 28 ValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaGluAlaGly 47

Db 307 GTTCTCCAGCTGCTCTTCTCTCCACCTACACGACG----- 345
Qy 48 ValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeu 67
Db 345 ----- 345
Qy 68 LeuGlnLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLys 87
Db 346 ----- 346
Qy 88 CysAspGluHisLysAlaSerSerProPheSerValAlaLysPheArgArgTyrAsp 107
Db 379 TTGAGCAACTACCCACATCCATTATGACCCCGGTTTCTTAATATACCTCCACGCT 438
Qy 108 CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu 127
Db 439 GTCGCCACGTCAGCAACCGCTTCAGCAACTAACCCAAATTCACAAATTCACAAATGTC 498
Qy 128 ProAlaLys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr 146
Db 499 ACATCCAGGCTTCAGCTCCACAGTCAAGAAATTCGAAATTAATTCATCAAGATTTGGT 558
Qy 147 PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGly 166
Db 559 ATCAGGCCAAGAGCTCTGACTTAACACCCCTCATCAACCACTTCATCAATAAT 618
Qy 167 LysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp----- 193
Db 619 ACRAAATTCATTCGCT-----CCGTCGAGCGGTTTCGAGTGGCAATAATAATGTGGC 669
Qy 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAla 198
Db 670 TCGACGATATCCACATCTCGAGAGGCTTAGAATCATCATCAACGTFACAGCTCTATTTCG 729
Qy 199 AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla 218
Db 730 AATCTAAACCGACCTACTCTCCAACTCCAAACCTTCTAGACCACAAACCCAGCTAGTT 789
Qy 219 AsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIle 238
Db 790 CGTGTCTCT-----ACAACTACAAATAATCGAAGC----- 819
Qy 239 ThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGlu 258
Db 820 -----TCAAAAGTAGCGCTCCGAAAGCCGTCGAGCACCCCAAACTTGTCTCT 867
Qy 259 ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu 275
Db 868 GTGAAGACTATTGGAGCAAAACAGACCCCGATACAGCCGCTGCTGTGTGTGGATG 927
Qy 276 ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer 290
Db 928 CTGAATTAAGTTATTTCAGTAGCAAAACCCATCTTCTCATCGAATAGCCACACCT 987
Qy 291 GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProVal 310
Db 988 ACGAGAAAGCGCGCGGCTGCTCAACAACTTTGCGAAATTCGTCGCCCGCAGTG 1047
Qy 311 ---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLys 329
Db 1048 AAAAGTGGCTGAAG-----COGCGCAGCTAGCTGGGAAGT 1086
Qy 330 ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
Db 1087 GCCACGCTATGTCGAAGCTTTGTACGCCAAAGTTCTTACCGTAAACCGGACGCCCA 1146
Qy 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
Db 1147 ATCATATCTCAACA-----GACTCG 1167
Qy 362 LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleLea 379
Db 1168 AACGATGCTCAAAAGACGAGTGAAGAGAGTCCGGATACGCTGTGATTTCAACAGCAGCTC- 1226

Qy 379 snAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAlaAsp 398
Db 1227 --GCCAAAGTTCATCATCGACGAGAGTTCCTTAAGCATGCTATCCACATCTTCCAGAGT 1284
Qy 399 ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle--- 417
Db 1285 TCAACGTCAAGAGAAAGTCTCCGTCATCAGACGATCTTACTCTTAAGCCCTCCATCGTG 1344
Qy 418 -----ProValSerAsnHisThrValGlyGluAsp 427
Db 1345 ACAGCTATCAGACAGCCGATAGCCGCAACACCGGTTTCTCCAAATATATTATCAACAGCT 1404
Qy 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db 1405 GTTGAGGAAAAACCAACACTGCGCAGTGAAGAGTGAAGAGCAGCGGAAAGATCCA 1464
Qy 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
Db 1465 CTTCCAGCTGTTCCGCCAGCTGACAC-----CAGCCACA 1500
Qy 461 ThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysVal 480
Db 1501 ATCGGAGTTGTTAGTCCCAATTATGACAT-----AAGAGTTG 1539
Qy 481 Thr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
Db 1540 ACAATATGACCCGCTGATATCTGAAAAACCAAGACCTGAAAGCTCAATCAATGAGCATC 1599
Qy 499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLysSerThrGln- 517
Db 1600 GACACG-----ACGACGTTCCACCGCTTCCACCTCTAAATCACTGTGTCCA 1647
Qy 518 -----ArgCysSerSerLysGlyLysThrA 526
Db 1648 CTTAAATGACTTCAATCCGACACCAACACCTGAGTGTCTCTTAAACAAGS----- 1703
Qy 526 IaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerT 546
Db 1704 -----AAAAATCATCGCTGTCAAG--TCGTGTGATGATGAGCAT 1744
Qy 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db 1745 CG-----TCGCGCTCTG 1756
Qy 566 LuAsnSerValLeuSerHis---SerAlaLysValSerPro-----AlaG 580
Db 1757 AGACTCTATTGCTGCTCATGCTCGGTCAGTCACTCCGCGCACAAAACTTCTGTGTA 1816
Qy 580 LuHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 600
Db 1817 ATCATTTGCTG-----GAGAGAGGATGGGAAGAATAAGACAT 1855
Qy 600 ysGlnLysLeuGluValThr-----ArgGluL 609
Db 1856 CAGATCCAGCGGTACACTCTGACCGCGGTGTGOGATGTGCGCAAAATGAGGAGA 1915
Qy 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db 1916 AGCTGAAGAATAACGATGATGACTCTGT-----CGAGCACGAAGCGCT 1960
Qy 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db 1961 ATCTTGACAACTTGGAGACAGTCTCTCTCTGCTGCTGGAATATCCGATACAAACGAGC 2020
Qy 645 leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
Db 2021 TCGACGACATATCCACGAGCAGATTTGTCCGAGTAGACATGGCAACAGTCGCTCCAAAC 2080
Qy 662 spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
Db 2081 AT---AGGACTATTCCCACTTTGTTGCGCATCCCACTCTTCTTCTCTCAAGCCCCGAG 2137

679	Qy	-----SerLeuAlaSerGlnSerThrGlnLysGluLeuLeuGlnGlyHisLeuAlaLeuThrT	697
2138	Db	TCGCCAGTCGGTCTCTCCACATCAGTCGATCTCGATCCGACAGAA-----	2184
697	Qy	hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuA	717
2185	Db	--CAGGAGATGTGTACAAACTCTCTCCAGTCCGACGCGCAACGACCAACGTGC-	2235
717	Qy	rgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis-----	734
2236	Db	-----GCCCTGCCACCTCAACTCTCGACACATTCGCTAA	2272
735	Qy	--AspAspGlnTyrIleAlaGluAlaProThrGluHisTyrGlyArgLysAspAlaLysL	754
2273	Db	GATCCCGGGATACTCATCTATTCTCCA-----CACTTATCA-----GTGTCACTGAT	2323
754	Qy	ysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC	772
2324	Db	AGGACAAATGTCTATGCACCTCAGACAGTAGTCGACGACCTTCTTCACAAAAACCAAGCT	2383
772	Qy	ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG	792
2384	Db	ATTGAGGCCAATT-----CAATCACTTGAT-----CGTAAATGCACTTCAAG	2428
792	Qy	lySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG	812
2429	Db	AGTTTCATCATCCCGACGACAGATGCGGCTCTCTGTGACCG-----	2472
812	Qy	luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA	832
2473	Db	-----AGACGGGTGCGGAACCTCGATCGAAATATGATT	2506
832	Qy	laSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrProLysG	850
2507	Db	CTTCAGATCTACTCGCGGTTCCGAGGTGGAAGCTCTACTGGTATCTATGGAGAG-	2565
850	Qy	luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP	870
2566	Db	-----ACGTTCCAACTGCACAGACTATCCGATGAAAAATCCCGGCACAT-----	2610
870	Qy	roAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA	890
2611	Db	-----TCTGCCAAAGTGGATGCCAACTA-----	2640
890	Qy	laHisAsnGlnTyrLysGlySerThrSerTyrGlySerAsnLeuAsnGlyLysI	910
2641	Db	-----TCACTGGCTAGCAGCAGCATATGCTCT-----CTCAATGAGAAGT	2683
910	Qy	leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu-	929
2684	Db	AC-----CAACATGCTATTCCGGACATGCGACGTGACTTGG	2719
930	Qy	-----ArgProHisProArgValGlyValLeuG	939
2720	Db	AGTGTACAGAACACTGCTCACTCAACCAAGAACAGGAACTATGGACATTGT	2779
939	Qy	lySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer----	957
2780	Db	TTGATCTTTTGACAAAGCTTAGAAAACTCACTCAACCAATTTGATCGATCCAACATGA	2839
958	Qy	-----GlyTyrLysLeuGlyValSerThrG	966
2840	Db	AGCCTGAGAGGCCAATACGATTACGGCAGGACATTGCTCAATTCAGGGATATTAGCAATC	2899
966	Qy	lyIleThrSerHisGlnMetAsnArgLysGlu-----	976
2900	Db	ATCTTGATCCAACTCACTAGCTCATGCTAACGAAGCGGTGTGAGCTTCTTCGTCAACCAT	2959
977	Qy	-----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysT	990
2960	Db	CTCTGGGAATCAGTTCGATCCCATCGATCGATCGATCTGCTCTCGAAGACGACGACG	3019
990	Qy	rpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsnS	1010

```

Db      3020 AGGAGAGATCACTGAGCTGCTGTTGGCAAGAAC----- 3054
Qy      1010 erilealaGlnSerTTPThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
Db      3055 -----AAGAGAGAGCTGGATCGCTCCCTCA-----CTCCCAAGTCA 3091
Qy      1030 alargGlnAspIleCysIleThrAsnLysasn-----ProAlaaspPheThrThrIles 1048
Db      3092 CCAAGAG-----AAGAACAGAAGACTACGACGAAGCACATATGCCATCAATTT 3139
Qy      1048 erAsnAspAsnGluTyrMetAsp 1055
Db      3140 CCGGATCTCAAGGAAGCTCTTGAC 3162

RESULT 4
US-09-845-917A-55
; Sequence 55, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845, 917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 4749
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-845-917A-55

Alignment Scores:
Pred. No.: 5,71e-10 Length: 4749
Score: 214.00 Matches: 243
Percent Similarity: 34.68% Conservative: 170
Best Local Similarity: 20.40% Mismatches: 414
Query Match: 3.87% Indels: 365
DB: 4 Gaps: 56

US-09-721-114-2 (1-1057) x US-09-845-917A-55 (1-4749)
Qy      14 ValGlyThrAsnCysMet-----LeuAlaArgGlyGlyThrGlyVala 27
Db      247 CTGGGCTCGACTGCTCGAAACTCACCAAAACCGATATCGACAGCGAACTTGGGTGCA 306
Qy      28 ValalaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaGluAlaGly 47
Db      307 GTTCTCCAGCTGCTCTTCTGCTCTCCACCTACAAGCAG----- 345
Qy      48 ValaspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeu 67
Db      345 ----- 345
Qy      68 LeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLys 87
Db      346 -----AAGCTTCGGCAACTGAAAGAAAGATCAGAGAAA 378
Qy      88 CysaspGluHisLysAlaSerSerSerProPheSerValalalysPheArgA-gTtpAsp 107
Db      379 TTGAGCAACTACCCACCATTCATTCACCCCGGGTTTCTAAATACCTCGCCACGT 438
Qy      108 CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu 127
Db      439 GTCCCAAGCTCAGCAACCGCTTCAGCAACTAACCCAAATTCCAACTTCCACAAATGCA 498
Qy      128 Proalalys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValargSerThr 146

```



```

Qy 772 ysGlyAlaGlnPheArgProGlyTleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db 2384 ATTCAGGCCAATTT-----CATTCACTTGAT---CGTAAATGCCACTTCAAG 2428
Qy 792 lySerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
Db 2429 AGTTCACATCCCGCAGCAGAGATGCGGCTCTCTTGAGCCG-----2472
Qy 812 luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
Db 2473 -----AGCGGGTGCCGAACTCGATTCGAAATATGATG 2506
Qy 832 laSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
Db 2507 CTTTCAGATCTACTCGCGCGTCCCGAGGTGGAAGCTCTACTGGTATCTATGGAGAG- 2565
Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
Db 2566 -----ACGTTCCAACTGCACAGACTATCCGATGAAATAATCCCGCAT-----2610
Qy 870 roAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA 890
Db 2611 -----TCGCCAAAGTGGATGGATCCCACTA-----2640
Qy 890 laHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysI 910
Db 2641 -----TCACTGCTAGCAGCAGCATATGATCT---CTCAATGAGAAGT 2683
Qy 910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
Db 2684 AC-----GAACATGCTATTCCGGACATGGCAGCTGACTTGG 2719
Qy 930 -----ArgProHisProArgValGlyValLeuG 939
Db 2720 AGTGTTACAGAACACTGTCGACTCACTAACCAAGAACAGGAGAACTATGGAGCATTGT 2779
Qy 939 lySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer-----957
Db 2780 TTGATCTTTTGGACAAAGCTTAGAAAACTCACTCAACACATTCGATCGCAACTTGA 2839
Qy 958 -----GlyTyrLysLeuGlyValSerThrG 966
Db 2840 AGCCTGAAGGCGCATACGATTACGGCAGCAGCATTGCTCATTTGAGGGATATTAGCATC 2899
Qy 966 lyIleThrSerHisGlnMetAsnArgLysGlu-----976
Db 2900 ATCTTGATCCAACTCAGCTCATGCTAACGAAGCGCTGTGAGCTTCTTCGTCAACCAT 2959
Qy 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaIleT 990
Db 2960 CTCTGGAATCAGTTGCATCCCATCGATCATCGATGTCGTCGCGAAAGCAGCAAGC 3019
Qy 990 rpAsnAlaLeuGlnLeuGlySerValSerSerAlaAspPheLeuSerAlaArgAsnS 1010
Db 3020 AGGAGAAGATCAGCTGAGCTCGTTTGGCAAGAAC-----3054
Qy 1010 erIleAlaGlnSerTyrThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
Db 3055 -----AAGAAAGCTGGATCCGCTCTCA-----CTCTCCAAGTTCA 3091
Qy 1030 alArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrIleS 1048
Db 3092 CCAAGAG-----AAGNACAAAGACTACGACGAGCAGCATATGCCCATCAATT 3139
Qy 1048 erAsnAspAsnGluTyrMetAsp 1055
Db 3140 CCGGATCTCAGGAACCTCTTGAC 3162

```

RESULT 5

US-09-845-917A-52

; Sequence 52, Application US/09845917A

; Patent No. 6653529

; GENERAL INFORMATION:

```

; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; US-09-845-917A-52

```

Alignment Scores:

```

Pred. No.: 6,26e-10 Length: 5009
Score: 214.00 Matches: 243
Percent Similarity: 34.68% Conservative: 170
Best Local Similarity: 20.40% Mismatches: 414
Query Match: 3.87% Indels: 365
DB: 4 Gaps: 56

```

US-09-721-114-2 (1-1057) x US-09-845-917A-52 (1-5009)

```

Qy 14 ValGlyThrAspCysMet-----LeuAlaArgGlyGlyThrGlyAla 27
Db 247 CTGGGTCTCGACTGCTCGAAACTCACCACAAACCGATATCGACAGCGAACTTGGGTGCA 306
Qy 28 ValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaGluAlaGly 47
Db 307 GTTCTCCAGTGTCTTCTCTCTCCACTACAGCAG-----345
Qy 48 ValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaIleu 67
Db 345 -----345
Qy 68 LeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLys 87
Db 346 -----AAGCTTCGCACTGAAAGAGATCAGAGAAA 378
Qy 88 CysAspGluHisLysAlaSerSerProPheSerValAlaLysPheArgTrpAsp 107
Db 379 TTGGACCAACTACCCATCCATTATGCCCGGGTTCTAAATTTACCTCGCCACGT 438
Qy 108 CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu 127
Db 439 GTGCCACGTCAGCAACCGCTTCAGCACTAACCCAAATCCCACTTCCACAAATGTCA 498
Qy 128 ProAlaLys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr 146
Db 499 ACATCCAGGCTTCAGACTCCAGTCCAGATATCGAAATTTGATTCATCAAGATTTGT 558
Qy 147 PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerGlnGly 166
Db 559 ATCAAGCCAAAGACGCTCGACTTAAACCCCTCATCATCAACCACTTCATCAATAAT 618
Qy 167 LysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp-----183
Db 619 ACAAAATTCATTCGCT-----CGTCGAGCGGTTCGATGGCAATAATAATTTGGC 669
Qy 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaGluAla 198
Db 670 TCGACGATATCCACTCTCGAAGAGCTTAGAATCATCATCAACGTCAGCTCTATTTCG 729
Qy 199 AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla 218
Db 730 AATCTAAACCGACCTACCTCCCACTCCAAAACCTTCTAGACCAACACCCAGCTAGTT 789
Qy 219 AsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIle 238

```

Db 790 CGTGTTCCT-----ACAACTACAAAATCGAAGC----- 819
 Qy 239 ThrTrpHisIleGluValAsnGlyAlaAspGlnProSerThrProLysLeuSerGlu 258
 Db 820 -----TCAAAGCTAGCGGTCCGAAAGCGGTGAGCACCAAACTTGTCTCT 867
 Qy 259 ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu 275
 Db 868 GTGAAGACTATTGAGCAAAACAGAGCCCGGATAAAGCGGTGGTGGTGGTGGTGAATG 927
 Qy 276 ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer 290
 Db 928 CTGAATTAAGTATTTCAGTACCAAAACCCATCTTCCTCAATCGAATAGCCCAACCT 987
 Qy 291 GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProVal 310
 Db 988 ACAGAAAGCGCGCGGTGCTCAACCAAACTTTGTGAAATCGCTCCCGCAGTG 1047
 Qy 311 ---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLys 329
 Db 1048 AAAAGTGGCTGAAG-----CGCCGACCAAGTAAAGCTGGAGT 1086
 Qy 330 ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
 Db 1087 GCCACGTCTATGTCGAAGCTTTGTAGCCCAAAAGTTTCTTACCGTAAACAGGCGCCCA 1146
 Qy 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGluAlaAspValSerAspAla 361
 Db 1147 ATCTATCTCAACA-----GACTCG 1167
 Qy 362 LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleAla 379
 Db 1168 AAACGATGCTCAAAGCAGTGAAGAGTCCGATACGCTGGATTCAACAGCAGCTC- 1226
 Qy 379 scAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAsp 398
 Db 1227 --GCCAAGCTCATCATCGAGGAGTTCCTTAAGCATGATCCACATCTTCCAGAGT 1284
 Qy 399 ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle--- 417
 Db 1285 TCAACGTGACAGCAAAAGTCTCGTCATCAGCAGATCTTACTTTAAAGCTCCATCGTG 1344
 Qy 418 -----ProValSerAsnHisThrValGlyGluAsp 427
 Db 1345 ACAGCTATCAGACAGCCGATAGCCGCAACACCGGTTTCTCCAAATATTATCAACAGCT 1404
 Qy 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
 Db 1405 GTTGAAGAAAACCAACACTGGCAGTGAAGAGTGAAGAGCACGCAAGAAAGATCCA 1464
 Qy 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
 Db 1465 CTCCAGCTGTTCCGCAAGTGACACC-----CAGCCCAACA 1500
 Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
 Db 1501 ATCGGAGTTGTAGTCAATATGCGCAT-----AAGAGTTG 1539
 Qy 481 Thr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
 Db 1540 ACAATGACCCGTGATATCTGAAACCAAGACCACTGAAAGCTCCAAATCAATGAGCATC 1599
 Qy 499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluIleSerThrGln- 517
 Db 1600 GACACG-----ACGGACGTTCCACCGCTTCCACCTCTAAATACATGTTGTCCA 1647
 Qy 518 -----ArgCysSerSerLysGlyLysThr 526
 Db 1648 CTTAAATATGACTTCAATTCGACCAACCAACCAAGTACATGATGTTCTTCAAAACAAG- 1703
 Qy 526 laGlyLeuSerLysGlyLysThrHisSerAlaAsnSerThrLysTyrGlyGlyGluSerT 546
 Db 1704 -----AAAAATCATCGCTGTCAAG--TCGTTTGGATATGAGCAGT 1744

Qy 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
 Db 1745 CG-----TCCGCGTCTG 1756
 Qy 566 luAsnSerValLeuSerHis---SerAlaLysValSerPro-----AlaG 580
 Db 1757 ARGATCCATTGTGCTCATCGTCCGCTCAGGTACTCCGCGCAAAAACCTTCTGGTA 1816
 Qy 580 luHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 600
 Db 1817 ATCATTCGCTG-----GAGAGAAGGATGGGAAAGATAGACAT 1855
 Qy 600 ysGlnLysLeuGluValThr-----ArgGluL 609
 Db 1856 CAGAAATCCAGCGGTACACCTCTGACCGCGGTGTTCGATGTGCGCCAAATATGAGGGA 1915
 Qy 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
 Db 1916 AGCTGAAGATACGATCATGACTCGT-----CGAGCACAGAAGCGGT 1960
 Qy 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
 Db 1961 ATCCTGACAACCTCGACAGCAGTTCTCTCTGCTGCTGGAATATCCGATACACGAGC 2020
 Qy 645 leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
 Db 2021 TCGACGACATATCCAGCGACGATTTGTCGGAGTAGACATGCGCAACAGTCCGCTCCAAC 2080
 Qy 662 spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
 Db 2081 AT---AGCGACTATTCCCACTTGTTCGCCATCCACAGCTTCTTCTCTCAAAAGCCCCGAG 2137
 Qy 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThr 697
 Db 2138 TCCCGAGTCGCTCCACATCGATTCGATTCGATTCGAGCAGAA----- 2184
 Qy 697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuA 717
 Db 2185 --CAGGAGAATGTGTACAAACTTCTGTCCAGTCGCAACGAGCAACAGCTGCGC----- 2235
 Qy 717 rgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
 Db 2236 -----GCCGCTGCCACTCCACTTCGCAACCTTCGCAACATTCGCTAA 2272
 Qy 735 --AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysL 754
 Db 2273 GATCCCGGGATCTCATCTCTATTCTCA-----CACTTATCA---GTGTCAGCTGATA 2323
 Qy 754 ysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC 772
 Db 2324 AGGACACAATCTCTATGCACCTCAGACACTAGTCGACGACCTTCTTCAAAAACCAAGCT 2383
 Qy 772 ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
 Db 2384 ATTCAGGCCAATTT-----CATTCACCTGAT---CGTAAATGCCACCTTCAAG 2428
 Qy 792 lysSerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
 Db 2429 AGTTACATATCCACCGACGACAGAAATGCGGTCTCTTTCGAGCCCG----- 2472
 Qy 812 luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
 Db 2473 -----AGACGGGTGCCGAACCTCGATGTCGAAATATGATT 2506
 Qy 832 laSerLysLeuCysAspArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
 Db 2507 CTTCAGGATCTACTCGCGCGGTTCCCGAGGTGGAAGCTCTACTGTTATCTATCGAGAG- 2565
 Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
 Db 2566 -----ACGTTCAACTGCACAGACTATCCGATGAAATAATCCCGCCACAT----- 2610

Db 1048 AAAAGTGGCTGAAG-----CCGCCGACCAGTAAGCTGGGAAGT 1086
Qy 330 ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
Db 1087 GCCACGTCATCTCGAAGCTTTGTACGCCAAAAGTTTCTACCGTAAACAGCAGCCGCCA 1146
Qy 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
Db 1147 ATCATATCTCAACAA-----GACTCG 1167
Qy 362 LysLeuLysArgArgLysProLysLysValArgLys-----LeuSerGluLysLys 379
Db 1168 AAACGATGCTCAAGAGCAGTGAAGAGAGTCCGATCGCTGGATTCAACAGCAGCTC- 1226
Qy 379 snAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAlaAsp 398
Db 1227 --GCCAACGTCATCATCGCAGGAAGTTCCTTAAGCATGTCATCCACATCTTCCAAAGT 1284
Qy 399 ProCysGluAspArgSerThrIleProValProMetGluValSerMetAspIle--- 417
Db 1285 TCAACGTCAGACGAAAGTCTCGTTCATCAGACGATCTTACTCTTAACGCTCCATCGT 1344
Qy 418 -----ProValSerAsnHisThrValGlyGluAsp 427
Db 1345 ACAGCTATCAGACGCGCATAGCCGCAACACCGGTTCTTCCAAATATTATCAACAGCCT 1404
Qy 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db 1405 GTTGAGGAAAAACCAACTGCGAGTGAAGAGTGAAGAGCAGCAGCAAGAAAGATCCA 1464
Qy 441 SerAspValValAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
Db 1465 CTTCCAGCTGTTCGCCAGCGTACACC-----CAGCCAA 1500
Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db 1501 ATCGGAGTTGTAGTCCAAATTATGGCAGAT-----AAGANGTTG 1539
Qy 481 Thr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
Db 1540 ACAAATGACCCCGTGATATCTCAAAACCCAGAACCTGAAAGCTCAATCAATGAGCATC 1599
Qy 499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLysThrGln- 517
Db 1600 GACAGC-----ACGACGCTTCCACGCTTCCACCTCTAAATTCAGTTGTTCCA 1647
Qy 518 -----ArgCysSerSerLysGlyLysThr 526
Db 1648 CTTAAATGACTTCAATCCGACCAACCAACGATAGATGTTCTTTCTAAACAGG---- 1703
Qy 526 laGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerT 546
Db 1704 -----AAAAATCATCATCGCTCTGCAAG--TCGTTTGGATATGACGAGT 1744
Qy 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db 1745 CG-----TCGCGGTCTG 1756
Qy 566 luAsnSerValLeuSerHis---SerAlaLysValSerPro-----AlaG 580
Db 1757 AAGACTCCATTGTGGTCTATGCTCGCTCGGTCAGTCCGCGCAAAACCTTCTGTTGA 1816
Qy 580 luHisAspIleGlnIleMetSerAspLeuHisGluLysSerLeuProLysLysLysL 600
Db 1817 ATCATTCGCTG-----GAGAGAGGATGGGAAGAATAAGACAT 1855
Qy 600 ysglnLysLeuGluValThr-----ArgGluL 609
Db 1856 CAGATCCAGCGCTACACTCTGACGCGGTGTTGCGATGTGCGCCAAATGAGGAGAGA 1915
Qy 609 ysglnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db 1916 AGCTGAAGAAGATACGATGATGACTCTG-----CGACGACAGACGCGCT 1960

Qy 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db 1961 ATCTGTGAACACTTCGAAGACAGTTCCTCTTGTGCTGGATATCGATAACACAGC 2020
Qy 645 leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
Db 2021 TCGACGACATATCCACGAGATTTGTCGGAGTAGACATCGCAACAGTCGCTCCAAAC 2080
Qy 662 spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
Db 2081 AT---AGCGACTATTCCCACTTTGTCGCCATCCCACTCTCTTCTTCCAAAGCCCCGAG 2137
Qy 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThr 697
Db 2138 TCCCGAGTCGCTCTCCACATTCAGTCATCTCGATCTCGACGAGAA----- 2184
Qy 697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnHisLeuA 717
Db 2185 --CAGGAGAATGTGTACAAACTTCTGTCCAGTGTCCGAGCGCAACGCTGCG- 2235
Qy 717 rgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis- 734
Db 2236 -----GCCGCTGCCACCTCAACCTTCGGACACATTCGCTAA 2272
Qy 735 --AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysL 754
Db 2273 GATCCCGGGATCTCATCTATTCTCCA-----CACTTATCA---GTGTCAGCTGATA 2323
Qy 754 ysgLysThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC 772
Db 2324 AGGACACATGTCTATGCTACACTCAGACTAGTCAGACCTTCTTCAAAAACCAAGCT 2383
Qy 772 ysgLysAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db 2384 ATTCAGGCCCAATTT-----CATTCACTTGAT---CGTAAATGCCACCTTCAAG 2428
Qy 792 lySerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
Db 2429 AGTTTCACATCCACGACGACAGATGCGGCTCTCTTGAGCCCG- 2472
Qy 812 luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
Db 2473 -----AGCGGGTCCGAACTCGATCGTCCGAATATGATT 2506
Qy 832 laSerLysLeuCysAspArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
Db 2507 CTTCAGCATCTACTCGCGGCTTCCGAGGTGGAAGCTCTACTGTATCTATGAGAG- 2565
Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
Db 2566 -----ACGTTCCAACTGCACGACTATCCGATGAATAATCCCCGACAT- 2610
Qy 870 roAsnTyrGlyThrSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA 890
Db 2611 -----TCTGCCAAAAGTCAGATGGATCCCACTA----- 2640
Qy 890 laHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysI 910
Db 2641 -----TCACTGGTAGCAGCAGCATATGATCT---CTCAATGAGAGT 2683
Qy 910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
Db 2684 AC-----GAACATGCTATTTCGGGACATGGCAGCTGACTGG 2719
Qy 930 -----ArgProHisProArgValGlyValLeuG 939
Db 2720 AGTGTTCAGAAACACTGTCGACTCACTAACCAAGAAACAGGAGAACTATGAGCATTGT 2779
Qy 939 lySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer---- 957
Db 2780 TTGATCTTTTGACAAAGCTTAGAAAACCTCACTCAACACATGATGCTCGATCCCACTGA 2839

Db 1348 TCAACGTGACGAGAAAGTCTCCGTCATCAGACGATCTTACTCTTAACGCCCTCCATCGT 1407
Qy 418 -----ProValSerAsnHisThrValGlyGluAsp 427
Db 1408 ACAGCTATCAGACAGCGGATAGCCGCAACACCGGTTCTCCAAATATTATCAACAGCCT 1467
Qy 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db 1468 GTTGAGGAAAAACCAACACTGCGCAGTGAAGAGTGAAAGCAGCAGCGGAAAAAGATCCA 1527
Qy 441 SerAspValValAspAspGlySerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
Db 1528 CTCCAGCTGTTCCGCCACGTGCACCC-----CAGCCCAACA 1563
Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db 1564 ATCGGAGTTGTAGTCCAAATTATGGCACAT-----AAGAAGTTG 1602
Qy 481 Thr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
Db 1603 ACMAATGACCCCGTGATATCTGMAAACCCAGAACCTGMAAAGCTCCMATCAATGAGCATC 1662
Qy 499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLysSerThrGln- 517
Db 1663 GACACG-----ACGGACGTTCCACCGCTTCCACCTCTAAATCACTGTGTCCA 1710
Qy 518 -----ArgCysSerSerLysGlyLysThr 526
Db 1711 CTTAAATGACTTCAATCCGACCAACCAACCACTAGATGTTCTCTTAAACAGG----- 1766
Qy 526 lagly-leuSerLysGlyLysThrHisSerAlaLysThrLysTyrGlyGlyGluSerT 546
Db 1767 -----AAAAATCACATCGCTGTCAAG--TCGTTTGGATATGACGACT 1807
Qy 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db 1808 CG-----TCGCGCTGTG 1819
Qy 566 luAsnSerValLeuSerHis--SerAlaLysValSerPro-----AlaG 580
Db 1820 AAGACTCCATGTGGTCTAGCTCGCTCGCTCAGTGACTCGCGCAGCAAAAATCTCTGGTA 1879
Qy 580 luHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 600
Db 1880 ATCATTCGCTG-----GAGAGAAGGTGGGAAAAAGATAGACAT 1918
Qy 600 ysGlnLysLeuGluValThr-----ArgGluL 609
Db 1919 CAGAAATCCAGCGGTACACTCTGACGCGGTGTCGGATGTCGCGCAAAATGAGGGAGA 1978
Qy 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db 1979 AGCTGAAAGAATACGATGACATGACTCGT-----CGAGCACAGAACGGCT 2023
Qy 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db 2024 ATCTTGACAACTTCGAGACAGTTCCTCTGTCGTCGTAATATCCGTAACACAGAGC 2083
Qy 645 leGlnSerLysThrAlaAsp-----AspAspCysValIleValAlaLysA 662
Db 2084 TCGACGACATATCCACGGAGATTGTTCGGAGTAGACATGGCAACAGTCGCTCCAAAC 2143
Qy 662 spGlySerAspTyrAlaSerVal---PheAspThrAsnSerGlnGlnLys----- 678
Db 2144 AT---AGCGACTATCCCACTTGTGTGGCCATCCACAGCTTCTCTCTCAAAGCCCGAG 2200
Qy 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrT 697
Db 2201 TCCCCAGTCGGTCTCCACATCAGTCGATCTCGATCTCGAGCAGAA----- 2247
Qy 697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuA 717
Db 2248 --CAGGAGATGTGTACAAACTTCTGTCCAGTGCAGGACGACCAACGTCGCG----- 2298

Qy 717 rgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
Db 2299 -----GCCGCTGCACCTCAACCTTCGACACACATTGCTAA 2335
Qy 735 --AspAspGlnTyrIleAlaGluAlaProThrGluHisTyrGlyArgLysAspAlaLysL 754
Db 2336 GATCCCGGATCTCATCTTATCTCCA-----CACTTATCA--GTGTCACTGATA 2386
Qy 754 ysLeuThrTrpGluGlnPheLysAlaThrArg-----AsnSerProAlaAlaThrC 772
Db 2387 AGGACACATGCTCTATGCATCTACAGACTAGTCAGCAGCTTCTTCAAAACCAAGC 2446
Qy 772 ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db 2447 ATTCAAGCCCAATT-----CACTCACTGAT--CGTAATGCACTTCAAG 2491
Qy 792 lysSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
Db 2492 AGTTCACTCCACCGAGCACAGAAATGGCGCTCTCTTGAGCCCG----- 2535
Qy 812 luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
Db 2536 -----AGCGGTGCGCACTCGATGTCGAAATATGATT 2569
Qy 832 laSerLysLeuCysAspArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
Db 2570 CTTCAAGATCTCTACTCGCGCGTTCCCGAGGTGGAAGCTCTACTGCTATCTATGAGAG 2628
Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
Db 2629 -----ACGTCCAACTGCACAGACTATCCGATGAAAAATCCCCCGCAT----- 2673
Qy 870 roAsnTyrGlyThrSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA 890
Db 2674 -----TCTGCCAAAAGTGAGATGGGATCCCACTA----- 2703
Qy 890 laHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysI 910
Db 2704 -----TCACTGGCTAGCACGACGATATGCGATCT--CTCAATGAGAAGT 2746
Qy 910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
Db 2747 AC-----GNACTGTATTTCGGGACATGGCAGCTGACTTGG 2782
Qy 930 -----ArgProHisProArgValGlyValLeuG 939
Db 2783 AGTGTTCACAGAACACTCTCGACTCACTAACCAAGAAACAGGAGAACTATGAGCATTTGT 2842
Qy 939 lysSerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer--- 957
Db 2843 TTGATCTTTTGGACAAAGCTTAGAAAACCTCCTCAACACATTCATCGATCCCACTGA 2902
Qy 958 -----GlyTyrLysLeuGlyValSerThrG 966
Db 2903 AGCCTGAAGAGCAATACGATTTCAGGCAGGACATTTGCTCAITTCAGGGATATTAGCAATC 2962
Qy 966 lyIleThrSerHisGlnMetAsnArgLysGlu----- 976
Db 2963 AFTCTTGATCCCACTCAGCTCATGCTAACGAAAGGCGCTGGTGAGCTTCTTGTGCAACCAT 3022
Qy 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysT 990
Db 3023 CTCCTGAATCAGTTCCATCCCATCGATCATCGATGTCATCGTCGCGAAAAGCAGCAAGC 3082
Qy 990 rpAsnAlaLeuGlnLeuGlySerValSerSerAlaAspPheLeuSerAlaArgAsnS 1010
Db 3083 AGGAGAAGATCAGCTTGAGCTGCTTTGGCAAGAAC----- 3117
Qy 1010 erIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
Db 3118 -----AAGAGAGCTGATTCGCTCCTCA-----CTCTCAAGTTCA 3154

1603 ACAATGACCCCGTGATATCTGAAACACAGAACCTGAAAGCTCAATCAATGAGCATC 1662
499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLeuSerThrGln- 517
1663 GACACG-----ACGAGCTTCCACCGCTTCCACCTCTAAATCATGTTTCCA 1710
518 -----ArgCysSerSerLysGlyLysThr 526
1711 CTTAAATGACTTCAATCCGACACACACCAACGATGCTCTCTTAAACAAAG----- 1766
526 laGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysThrGlyGlyGluSerT 546
1767 -----AAAATCATCGCTGTCAAG--TCGTTGATATGAGCAGT 1807
546 h-ArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
1808 CG-----TCGCGGTCTG 1819
566 luAsnSerValLeuSerHis--SerAlaLysValSerPro-----AlaG 580
1820 AGACTCCCAATGGGCTCATGGCTCGCTCGCTGACTCGCGACAAAACCTTCGTGTA 1879
580 luHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 600
1880 ATCATTCGCTG-----GAGAGAAGGATGGGAAAGAAATAGACAT 1918
600 ysGlnLysLeuGluValThr-----ArgGluL 609
1919 CAGAAATCCAGCGGCTACACTCTGACGCGGTGTCGAGTGGCGCCAAATAGAGGAGA 1978
609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
1979 AGTGAAGATACGATGACATGACTGCT-----CGAGCACAGACGGCT 2023
629 is-----GluA-glnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
2024 ATCTGCACAACTTCGAAGACAGTCTCTCTGCTGCTGAATATCCGATAACACAGC 2083
645 leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
2084 TCACACATATCCAGGAGATTTGTCGGAGTAGACATGCGACAGTCGCTCCAAAC 2143
662 spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
2144 AT---AGCGACTATCCCACTTGTTCGCCATCCACGCTCTTCTCTCTCAAAAGCCCGAG 2200
679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrT 697
2201 TCCCGAGTCGGTCTCCACATCATCGATCTCGATCTCGACGAGAA----- 2247
697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuA 717
2248 --CAGGAGAAATGTGTACAAACTTCTGTCCAGTCCGACGACGACCAACGCTGCG- 2298
717 rgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
2299 -----GCGCTGCCACCTCAACCTTCGGACACCAATCCGCTAA 2335
735 --AspAspGlnTyrIleAlaGluAlaProThrGluHisThrGlyArgLysAspAlaLysL 754
2336 GATCCCGGATATCATCTCTATCTCA-----CACTTATCA---GTGTCAGCTGATA 2386
754 ysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC 772
2387 AGGACACAATGTCTATGCACTACAGACTAGTCGACGACCTCTTCCAAAAACCAAGCT 2446
772 ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
2447 ATTGAGGCCAATTT-----CACTCATTTGAT---CGTAATGCCACTTTCAG 2491
792 lysSerSerAsnThrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
2492 AGTTCATCCACCGACAGCAATGGCGGCTCTCTTGAGCCCG----- 2535

812 luA-gAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
2536 -----AGACGGGTGCGCAACTCGATGTCGAAATATGATTT 2569
832 laSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
2570 CTTGAGATCTCTACTCGCGGCTTCCCGAGGTGGAAGCTCTACTGTTATCTATGAGAG- 2628
850 luSerMetProAlaThrHisIleLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
2629 -----ACGTTCCAACTGCACAGACTATCCGATGAAAAATCCCCCCACAT----- 2673
870 roAsnTyrGlyThr-SerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA 890
2674 -----TCGCCAAAGTGGATGGATCCCACTA----- 2703
890 laHisAsnGlnTyrLysGlySerThrSerThrSerThrGlySerAsnLeuAsnGlyLysI 910
2704 -----TCACTGGCTAGCACACACAGATATGGATCT---CTCAATGAGAAGT 2746
910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
2747 AC-----GAACATGCTATTTCGGGACATGSCACGTGACTTGG 2782
930 -----ArgProHisProArgValGlyValLeuG 939
2783 AGTGTTCACAGACACTGTCGACTCACTAACCAAGAACAGAGAACTATGGAGCATTTGT 2842
939 lysSerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer---- 957
2843 TTGATCTTTTGGACAAAGCTTAGAAAACTCACTCAACATTCATGATCGATCCCACTGA 2902
958 -----GlyTyrLysLeuGlyValSerThrG 966
2903 AGCCTGAAGAGCAATACGATTCAGGAGGACATGCTCATTTGAGGATATTAGCAATC 2962
966 lyIleThrSerHisGlnMetAsnArgLysGlu----- 976
2963 ATCTTGCACTCAACTCAGCTCATGCTAACGAAGCGCTGGTGAGCTTCTTCGTCAACCAT 3022
977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysT 990
3023 CTCGTGAATCAGTTCATCCATCGATCATCGATGTCATCGTTCGAAAAAGCAGCAAGC 3082
990 rpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsnS 1010
3083 AGGAGAAGATCAGCTTGAGCTCGTTTGGCAAGAAC----- 3117
1010 erIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
3118 -----AAGAGAGCTGGATCCGCTCTCA-----CTCTCAAGTTCA 3154
1030 alaArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrThrIleS 1048
3155 CCAAGAG-----AAGAACAAAGAACTAGACGAAGACATATGATGATCAATTT 3202
1048 erAsnAspAsnGluTyrMetAsp 1055
3203 CCGGATCTCAAGGAAGTCTTGAC 3225

RESULT 9
US-09-845-917A-30
Sequence 30, Application US/09845917A
Patent No. 653529
GENERAL INFORMATION:
APPLICANT: Bogaert, Thierry
APPLICANT: Vandekerckhove, Joel
TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
FILE REFERENCE: P/14-1

Best Local Similarity: 20.40% Mismatches: 414
Query Match: 3.87% Indels: 365
DS: 4 Gaps: 56

US-09-721-114-2 (1-1057) x US-09-845-917A-25 (1-10443)

QY 14 ValGlyThrAsnCysMet-----LeuAlaArgGlyThrGlyAla 27
DB 257 CTGGCTCGACTGCTCGAAACTCACCAAACCGATATCGACAGCGAAACTGGGTGCA 316
QY 28 ValAlaProValLeuGluThrAlaThProArgGlnAspAlaAlaGluAlaGly 47
DB 317 GTTCTCCAGTGTCTTCTGCTCTCCACCTCAAGCAG----- 355
QY 48 ValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeu 67
DB 355 ----- 355
QY 68 LeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLys 87
DB 356 -----AAGCTTCGGCAACTCGAAAAAGATCAGAGAAA 388
QY 88 CysAspGluHisLysAlaSerSerProPheSerValAlaLysPheArgArgTrpAsp 107
DB 389 TTGGAGCAACTACCCACATCCATTATGCCACCGCGGTTTCTAAATTACCTCGCCACGT 448
QY 108 CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu 127
DB 449 GTGCCACGTCACCAACCGCTTCAGCACTAACCCAAATCCAACTTCCACAAATGTCA 508
QY 128 ProAlaLys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr 146
DB 509 ACATCCAGGCTTCAGACTCCACAGTCAAGAATATCGAAAAATGATTTCATCAAAAGATTGGT 568
QY 147 PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerGlnGly 166
DB 569 ATCAAGCCAAAGACGTCTGAGCTTAACACACCCCTCATCATCAACCTTCATCAATAAT 628
QY 167 LysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp----- 183
DB 629 ACAAAATTCATTCGGT-----CGGTGAGCGGTTCCAGTGGCAATAATAATGTTGGC 679
QY 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAla 198
DB 680 TCAGCATATCCATCTCGAAGAGCTTAGAATCATCATCAACGTACAGCTATATTCTG 739
QY 199 AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla 218
DB 740 AATCTAAACCGACTTACCTCCCACTCCAAABACCTTCTAGACCACAAACCCAGTAGTT 799
QY 219 AsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIle 238
DB 800 CGTGTGCT-----ACAATCAAAAATCGGAAGC----- 829
QY 239 ThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGlu 258
DB 830 -----TCAAGTAGCGCTTCGAAGCGGTGAGCACCCCAAACTTGCTTCT 877
QY 259 ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu 275
DB 878 GTGAAGACTATTGGACAAAACAAGACCGCATTAACAGCGGTGGTGGTGGTGAATG 937
QY 276 ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer 290
DB 938 CTGAATTAAGATTATTCAGTACAAAACCCATCTTCTCATCGAATAGCCCAACACT 997
QY 291 GlyLysGluArgAspGlnValAlaGlnCysAsnLeuThrLysAspProLysProVal 310
DB 998 ACGAGAAAGCGCGCGGTGCTTCAACCAACAACTTTGTCGAAATCGCTGCCCGAGTG 1057
QY 311 ---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluValValLeuLys 329
DB 1058 AAAAGTGGCTGAAG-----CGCCGACCAAGTAGTGGGAAGT 1096

QY 330 ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
DB 1097 GCCACGCTCTATGTCGAAGCTTTGTACGCCAAAGTTTCTACGTAAGAAACGGACGCCCA 1156
QY 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
DB 1157 ATCATATCTCAACA-----GAATCG 1177
QY 362 LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleAla 379
DB 1178 AAACGATGCTCAAGAGCAGTGAAGAAGTCCGGATACGCTGGATTCAACAGCAGCTC- 1236
QY 379 snAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAsp 398
DB 1237 --GCCACGTATCATCGACGAGGTTCCTTAAGCATGCTTCCATCTTCCAGAGT 1294
QY 399 ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle---- 417
DB 1295 TCAAGCTCAGACGAAAGTCTCCGTCATCAGACGATCTTACTCTTAACGCTCCATCGTG 1354
QY 418 -----ProValSerAsnHisThrValGlyGluAsp 427
DB 1355 ACAGCTATCAGACGCGGATACCCGACACCGGTTCTCCAAATATTATCAACAGGCT 1414
QY 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
DB 1415 GTTGAGGAAACCAACACTGCGAGTGAAGAGTGAAAGACACAGCGAAAGAGTCCA 1474
QY 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
DB 1475 CCTCAGCTGTTCCGCCACGTGACAC-----CAGCCAACA 1510
QY 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
DB 1511 ATCGAGTGTGTAGTCCAAATTATGACAT-----AAGAGTGTG 1549
QY 481 Thr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
DB 1550 ACAAAATGACCGCTGATATCTGAAACACAGAACCTGAAAGAGCTCAATCAATGAGCATC 1609
QY 499 AspThrAsnMetHisLysThrAspVal---CysGlnHisValSerGluIleSerThrGln- 517
DB 1610 GACAG-----ACGAGCTTCCACCGCTTCCACCTCTAAATCATGTTTCTTCCA 1657
QY 518 -----ArgCysSerSerLysGlyLysThr 526
DB 1658 CTTAAATGACTTCAATCCGACACACCAACGATGATGTTCTTTTAAACAAGS- 1713
QY 526 IagLysLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSert 546
DB 1714 -----AAAAATCATCATCGCTGTCAAG--TCGTTTGGATATGAGCAGT 1754
QY 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
DB 1755 CG-----TCCGCTCTG 1766
QY 566 IuAsnSerValLeuSerHis---SerAlaLysValSerPro-----AlaG 580
DB 1767 AAGACTCCATTTGGCTCATGCTCGGCTCAGGTGACTCCGCCGACAAAAAATCTCTGTA 1826
QY 580 IuHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysL 600
DB 1827 ATCATTCCTG-----GAGAGAAGCATGGGAAGATATAGACAT 1865
QY 600 ysGlnLysLeuGluValThr-----ArgGluL 609
DB 1866 CAGAATCCAGCGGCTTACACTCTGACGCGGTTGCGATGTCGCCAAATAATAGAGGAGA 1925
QY 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
DB 1926 AGCTGAAGAATAACATGACATGACTCT-----CGAGCAGACAGACGCT 1970

Qy	48	ValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeu	67
Db	346	-----	346
Qy	68	LeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLys	87
Db	347	-----AAGCTTCGGCAACTGAAAAAGATCACAAGAA	379
Qy	88	CysAspGluHisLysAlaSerSerProPheSerValAlaLysPheArgArgTyrAsp	107
Db	380	TTGAGCAACTACCACATTCATTATGCACCCCGGTTTCTAAATATACCTCCGCCAGT	439
Qy	108	CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu	127
Db	440	GTCCGCCAGTCAGCAACCGCTTCAGCAACTAACCCAAATTCCACTTCCCAAAATGTCA	499
Qy	128	ProAlaLys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr	146
Db	500	ACATCCAGGCTTCAGACTCCACAGTCAGAAATATCGAAATATGATTCATCAAGATTGGT	559
Qy	147	PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGly	166
Db	560	ATCAAGCCAAAGAGCTCGGACTTAAACCCACCTCATCATCAACACCTTCATCAATAAT	619
Qy	167	LysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp-----	183
Db	620	ACAAATTCAATCCGT-----CCGTCGAGCCGTTCCGAGTGGCAATAATAATGTTGGC	670
Qy	184	-----SerLysCysAsnAlaProSerGlyLysAsnGlyLysAlaAlaGluAla	198
Db	671	TCGACGATATCCACATCTCGAAGAGCTTAGAATCATCATCAACGTACAGCTCTATTTCG	730
Qy	199	AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla	218
Db	731	AAATCTAAACCGACTTACTCCCACTCCAAACAACTCTAGACACACAAACCCAGCTAGT	790
Qy	219	AsnValSerGluAspAsnThrSerValaspValGlyAlaLeuProGluValProGlnIle	238
Db	791	CGTGTGTCT-----ACAACACAAAATCGGAGC-----	820
Qy	239	ThrTrpHisIleGluValAsnGlyValaAspGlnProProSerThrProLysLeuSerGlu	258
Db	821	-----TCAAGACTAGCCGCTCCGAAAGCGGTGACACCCCAAACTGCTCTCT	869
Qy	259	ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu	275
Db	869	GTGAAGACTATTGGAGCAAAACAGACCCGATACAGCGGTGGTGGTGGGGAATG	928
Qy	276	ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer	290
Db	929	CTGAATTAATAAGTATTTCAGTAGCAAAAACCCATCTTCCTCATCGAATACGCCACACCT	988
Qy	291	GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProVal	310
Db	989	ACGAGAAGCGCGCGGTGCCTCAACACAAACTTTGTGAAAATCGCTGCCCCAGTG	1048
Qy	311	---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLys	329
Db	1049	AAAAGTGGCTCAAG-----COGCGCACCACTAAGCTGGGAAGT	1087
Qy	330	ArgSerSerLysSerLys-----ArgLysThrAspLysLys	341
Db	1088	GCCACGCTATGTCGAGACTTTGTAGCCCAAAAGTTTCCTACCGTAAACGGAGCGCCCA	1147
Qy	342	LeuMetLysLysGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla	361
Db	1148	ATCATATCTCAACAA-----GACTCG	1168
Qy	362	LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleIle	379
Db	1169	AAACGATGCTCAAAAGACGAGTGAAGAAGAGTCCGGATACGCTGGATTCAACAGCAGCTC-	1227

379	Qy	snlaaenGlnValGluuaspSerArg-SerAspGluValHisArgGluasnAlaAasp	398
1228	Db	--GCCAACGTCATCATCGCGGAAGTCTCCCTAAGCATGATCCACATCTTCCAGAGT	1285
399	Qy	ProCysGluuaspAspArgSerThrIleProValProMetGluValSerMetAspIle---	417
1286	Db	TCAACGTCAGACGAAAGTCTCGTCATCAGACGATCTTACTCTTAACGCGCTCCATCGTG	1345
418	Qy	-----ProValSerAsnHisThrValGlyGluuasp	427
1346	Db	ACAGCTATCAGACAGCGCGATAGCGCAACACCGGTTTCTCCAATATATTCAACAAGCCT	1405
428	Qy	-----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr	440
1406	Db	GTTGAGAAACCAACACCTCGCAGTCGAAGAGTGAAGACACAGCGAAGAAAGATCCA	1465
441	Qy	SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg	460
1466	Db	CCTCCAGCTGTTCGCCACGCTGACCC-----CAGCCCAACA	1501
461	Qy	ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal	480
1502	Db	ATCGAGTTGTAGTCCAAATTATGGCCAT-----AAGAAGTTG	1540
481	Qy	Thr-----ProThrAlaSerThrGlnHisAspAspGluasnAspThrGluAsnGlyLeu	498
1541	Db	ACAAATGACCCCGTGATATCTGAAAACACGAACTGTGAAAAGCTTCCAATCAATGAGCATC	1600
499	Qy	AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluIleSerThrGln-	517
1601	Db	GACACG-----ACGGACGTTCACCGCTCCACCTCTCAAATCAGTTGTGCCA	1648
518	Qy	-----ArgCysSerSerLysGlyLysThrA	526
1649	Db	CTTAAATGACTTCAATCCGACACACCACCACTGATGTTCTTTAAACAAGG----	1704
526	Qy	laGlyLeuSerLysGlyLysThrHisSerAlaIaSerThrLysTyrGlyGlyGluSerT	546
1705	Db	-----AAAATCACAATCCCTGCTCAAG--TCGTTTGGATATAGACAGT	1745
546	Qy	hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluuaspGlnCysGlnMetGluThrG	566
1746	Db	CG-----TCGCGCTCTG	1757
566	Qy	luAsnSerValLeuSerHis--SerAlaLysValSerPro-----AlaG	580
1758	Db	AAGACTCCATTGGCTCATCGCTCGGCTCAGGTGACTCGCGCAGCAAAAACTTCTGGTA	1817
580	Qy	luHisAspIleGlnIleMetSerAspLeuHisGlnGlnSerLeuProLysLysLysL	600
1818	Db	ATCATTCGCTG-----CAGAGAAGGATGGGAAGAATAAGACAT	1856
600	Qy	ysGlnLysLeuGluValThr-----ArgGluL	609
1857	Db	CAGAAATCCGCGCTACACTCTGACCGCGTGTGGATGTGCGCCAAATGAGGGAGA	1916
609	Qy	ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH	629
1917	Db	AGCTGAAAGAATACGATGATGACTCGT-----CCAGCACACAGACGGCT	1961
629	Qy	is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI	645
1962	Db	ATCCTGCAACTTGAAGACAGTTCCTCTTGTGCTCTGGATATCCGATACACACGACG	2021
645	Qy	leGlnSerLysThrThrAlaAasp-----AspAspCysValIleValAlaAlaLysA	662
2022	Db	TOGACGACATATCCACGGACGATTTGTCGGAGTACAGATGCGACACAGTCGCTCCAAC	2081
662	Qy	spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys-----	678
2082	Db	AT---ACGCGACTATTCCACTTTGTGCGCATCCCACTGCTTCTTCTCCTCAAGCCCGCAG	2138
679	Qy	-----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrT	697

Db 2139 TCCCAAGTGGTCTCCATCATGATCTCGATCTCGACGAGAA----- 2185
Qy 697 hGlnGlnSerProHisProGlnAsnPhelGlnSerThrGlnGlnGlnThrHisLeuA 717
Db 2186 --CAGGAGATGTGTACAAACTCTCTCCGACGCGAAGCAGCCAACTGGC----- 2236
Qy 717 rGmetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
Db 2237 -----GCCGCTGCCACCTCAACCTTCGGACCAACATCGCTTAA 2273
Qy 735 --AspAspGlnTyrlleAlaGluAlaProThrGluHisThrGlyArgLysAspAlaLysL 754
Db 2274 GATCCCGGGATACTCATCTATCTCCCA-----CACTTATCA---GTGTCAGCTGATA 2324
Qy 754 yLeuThrThrGluGlnPheLysAlaThrArg-----AsnSerProAlaAlaThrC 772
Db 2325 AGACACATGCTATGCTACACACACTAGTCGACGACCTTCTTACAAAACCAAGCT 2384
Qy 772 yeGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db 2385 AITCAGGCCAAATTT-----CACTCACTGAT---CGTAAATGCCACCTTCAAG 2429
Qy 792 lySerSerSerAsnTyrlaSerArgGlnProValIleAlaProLeuAspArgTyrlaAlaG 812
Db 2430 AGTTCATCCACCGACGACAGATCGCGCTCTCTTGAGCCCG----- 2473
Qy 812 luArgAlaValaAsnGlnValHisAlaArgAsnPhelProSerThrIleAlaThrMetGluA 832
Db 2474 -----AGACGGTGGCGAAGCTCGATGTCGAAATATGATT 2507
Qy 832 laSerlyLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrlProLysG 850
Db 2508 CTTCAAGATCTTACTCGCGCGCTCCGAGGTGGAGCTCTACTGTATCTATGAGAG- 2566
Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
Db 2567 -----ACGTTCCAACTGCACAGACTATCCGATGAAAATCCCGGCACAT---- 2611
Qy 870 roAsnTyrlThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyra 890
Db 2612 -----TCTGCCAAAGTGAGATGGATCCCAACTA----- 2641
Qy 890 laHisAsnGlnTyrlLysGlySerThrSerThrSerThrSerThrSerThrSerThrSer 910
Db 2642 -----TCACGTGGTAGCAGCAGCAGCATATGATCT---CTCAATGAGAGT 2684
Qy 910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
Db 2685 AC-----GAACATGCTATTTCGGACATGGCAGCTGACTTGG 2720
Qy 930 -----ArgProHisProArgValGlyValLeuG 939
Db 2721 AGTGTACAGAACACTGTCGACTCTAACCAGAAACAGAGAGAACTATGAGCAATTGT 2780
Qy 939 lySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer----- 957
Db 2781 TTGATCTTTTGAGCAAAAGCTTAGAAACTCACTCAACACATTGATCGATCCAACTTGA 2840
Qy 958 -----GlyTyrlLysLeuGlyValSerThrG 966
Db 2841 AGCCTGAAGAGGCAATACGATTTCAGCAGCAGCATTCCTCATTTGAGGGATATTAGCAATC 2900
Qy 966 lyIleThrSerHisGlnMetAsnArgLysGlu----- 976
Db 2901 ATCTTCATCCAACTCAGCTCATGCTAAACGAGGCGCTGTCGAGCTTCTTCGTCAACCAT 2960
Qy 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysT 990
Db 2961 CTTGGAATCAGTTGTCATCCATTCGATCATGCTATGCTGCTGCTGCGAAGACGCAAGC 3020
Qy 990 rPAsnAlaLeuGlnLeuGlySerValSerSerAlaAspPheLeuSerAlaArgAsnS 1010

Db 3021 AGGAGAAGATCAGCTTGCAGCTGGTTTGGCAAGAAC----- 3055
Qy 1010 erIleAlaGlnSerThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
Db 3056 -----AAGAAAGCTGGATCGCTCCTCA-----CTTCCAGTTTCA 3092
Qy 1030 alArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrThriles 1048
Db 3093 CCAAGAAAG-----AAGAACCAAGAACTACGACGAGGACATATGCCATCAATTT 3140
Qy 1048 erAsnAspAsnGluTyrlMetAsp 1055
Db 3141 CCGATCTCAAGGAACTCTTGAC 3163
RESULT 12
US-09-845-917A-26
; Sequence 26, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 7474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificial
; OTHER INFORMATION: plasmid
US-09-845-917A-26
Alignment Scores:
Pred. No.: 2,46e-09 Length: 7474
Score: 211.00 Matches: 254
Percent Similarity: 34.74% Conservative: 174
Best Local Similarity: 20.62% Mismatches: 415
Query Match: 3.82% Indels: 390
DB: 4 Gaps: 62
US-09-721-114-2 (1-1057) x US-09-845-917A-26 (1-7474)
Qy 8 GlnGluGlyAlaArgValValGlyThrAsnCysMetLeuAlaArgGlyGlyThrGlyAla 27
Db 347 AAGCAAGCGAAGAAAGCGAAGGAGCGGC-----GCTAGG---GCGCTGCCAAGT 394
Qy 28 ValalaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaGluAlaGly 47
Db 395 GTACGGTTCAGCTGCGCGTAACCAACCAACACCC-----GCCGCG 433
Qy 48 ValAspGluProAlaGln-----HisGlnCys---Glu 57
Db 434 CTTAATCGCGCTACAGGCGCGTCCCATTCGCAATTCAGCTGCGCAACTGTTGGCAA 493
Qy 58 HisPheSerIleArgGlyTyrlValalaLeuLeuGlnLysAspProLysPheCysSer 77
Db 494 GGGCGATCGGTGCGCGCC---TCTTCGCTATTACGCCAGCTGGCGAAGGGGGATGTGCT 550
Qy 78 LeuSerArgIle-----PheHisAspGlnLysLysCysAspGluHis 91
Db 551 GCAAGCGATTAACTGGTTAAGCCAGGGTTTCCCGAGTCCAGCTGTTGTTAAACGACG 610
Qy 92 LysAlaSerSerSerProPheSerValalaLysPheArgArgTrpAspCysSerLysCys 111
Db 611 GCCAGTGAAGCGCGGTAAATAC----- 631

Db 2487 -----GCCGCTGCCACCTCAACCTCCGACCAACATTCGCTA 2522
QY 735 ---AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLys 753
Db 2523 AGATCCCGGGATACCTCATCTCTCTCA-----CACTTATCA-----GTGTCACTGAT 2573
QY 754 LysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThr 771
Db 2574 AAGGACACATCTCATGCACTACAGACTAGTCAGCAGCTCTCTCAAAAACCAAGC 2633
QY 772 CysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMet 791
Db 2634 TATTCAGGCGCAATTT-----CACTCATGTAT---CGTAATGCCACCTTCAA 2678
QY 792 GlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAla 811
Db 2679 GAGTTCACATCCAGCAGCAGACAGATGGCGGCTCTCTTGAGCGCG----- 2723
QY 812 GluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGlu 831
Db 2724 -----AGACGGGTGCCGACACTCGATGTGCGAAATATGAT 2756
QY 832 AlaSerLysLeuCysAspArgAsnAlaGly-----GlnValValLeuTyrProLys 849
Db 2757 TCTTCAGGATCTACTCGCGGTTCCCGAGTGGNAGCTCTACTGGTATCTATCGAGAG 2816
QY 850 GluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhe 869
Db 2817 -----ACGTTCACATGCACAGACTATCGGATGAAATAATCCCGCCACAT--- 2861
QY 870 ProAsnTyrGlyThrSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyr 889
Db 2862 -----TCTCCAAAGTGGATGGGATCCCACTA----- 2891
QY 890 AlaHisAsnGlnTyrLysGlySerThrSerThrSerThrGlySerAsnLeuAsnGlyLys 909
Db 2892 -----TCACTGGCTAGCAGCAGCATATGATCTCT---CTCAATCAGAG 2933
QY 910 IleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu 929
Db 2934 TAC-----GAACATGCTTATTCGGACATCGACATG 2969
QY 930 -----ArgProHisProArgValGlyValLeu 938
Db 2970 GAGTGTACAGAACACTGCTGACTCACTAACAGAAACAGAGAACTATGAGCATTTG 3029
QY 939 GlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer--- 957
Db 3030 TTTGATCTTTTTCAGCAAAAGCTTAGAAAACTCACTCAACACATTCGATCGCAACTTG 3089
QY 958 -----GlyTyrLysLeuGlyValSerThr 965
Db 3090 AAGCCTGAAGAGCAATACATTCAGCGAGGACATTTGCTCATTTGAGGATATTAGCAAT 3149
QY 966 GlyIleThrSerHisGlnMetAsnArgLysGlu----- 976
Db 3150 CATCTTGCATCCAACTCAGCTCATGCTAACGAGGCGCTGGTGGAGCTTCTTGTCAACCA 3209
QY 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLys 989
Db 3210 TCTCTGGAATCAGTTGATCCCATCCATCATTCGATGTCATTCGTCGAAAGCAGCAAG 3269
QY 990 TrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsn 1009
Db 3270 CAGGAGAAGATCAGCTTGAGCTCGTTTGGCAAGAAC----- 3305
QY 1010 SerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPhe 1029
Db 3306 -----AAGAAGAGCTGGATCGCTCTCTCA-----CTCTCAAGTTTC 3341
QY 1030 ValArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrThrIle 1047
Db 3342 ACCAAGAG-----AAGAACAAGAACTACGAGCAAGACACATATGCTCAAT 3389

QY 1048 SerAsnAspAsnGlnTyrMetAsp 1055
Db 3390 TCCGGATCTCAAGCACTCTTGAC 3413
RESULT 13
US-09-845-917A-49
; Sequence 49, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandeckerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 4217
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-845-917A-49
Alignment Scores:
Pred. No.: 3,51e-09 Length: 4217
Score: 205.00 Matches: 232
Percent Similarity: 34.86% Conservative: 156
Best Local Similarity: 20.84% Mismatches: 379
Query Match: 3.71% Indels: 347
DB: 4 Gaps: 54

US-09-721-114-2 (1-1057) x US-09-845-917A-49 (1-4217)

QY 110 LysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAla 129
Db 6 AAGAAATGGAGCAACTACCCACATCCATTATGCGCCGGGTTTCTAAATTACCTCG 65
QY 130 LysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPhe--- 147
Db 66 CCAGTGTGCGCACGTCCAGCAACCGCTTCAGCAACTAACCCAAATTCCTCACTTCCCAA 125
QY 147 ----- 147
Db 126 ATGTCAACATCCAGCTTCAGACTCCACAGTCAAGAAATATCGAAATTTGATTCAATAAG 185
QY 148 -----ValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSer 164
Db 186 ATTGTATCAAGCCCAAGACGCTCTGGACTTAAACCCACCTCATCATCAACCACTTCATCA 245
QY 165 GlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp--- 183
Db 246 AATATACAAATTCATTCCTGCT-----CCGTGCGCCGCTTCGAGTGGCAATAATAAT 296
QY 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAla 196
Db 297 GTTGGCTCGACGATATCCACATCTCGAAGAGCTTAGAATCATCATCAACGTACAGCTCT 356
QY 197 GluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspVal 216
Db 357 ATTTGGAATCTTAACCGACCTACCTCCCACTCCAAAACCTCTTAGACCAACCAACCCAG 416
QY 217 AlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValPro 236
Db 417 CTAGTTCTGTGTGCT-----ACAACATAAAATCGGAAGC----- 452
QY 237 GlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeu 256
Db 453 -----TCAAGCTAGCCCTCGAAGAGCGGTGAGCCAGCCCAAACTT 494

Qy	257	SerGluValVal	-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGlu	273
Db	495	GCTTCTCTGAAGACTATTGGACGAAAAACAAGACCGCGTGGTGGTGGTGGT	554	
Qy	274	ThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro	-----AsnPro	288
Db	555	GGAAATGCTGAATTAAGATTATTACAGTACGAAACCCATCTTCTCTCATCGATATAGCCCA	614	
Qy	289	MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys	308	
Db	615	CAACCTACGAGAAAGGGGGCGGCGCTCCACACACAACTTTGTGCAAAATCGCTGCC	674	
Qy	309	ProVal---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValVal	327	
Db	675	CCAGTGAAGATGCGCTGAG-----CCGCGCACCATGATGCTG	713	
Qy	328	LeuLysArgSerSerLysSerLys-----ArgLysThrAsp	339	
Db	714	GGAAGTCCACGCTCTATGTCGAAGCTTTGTAGCGCAAAAGTTTCTTACCGTAAACCGGAC	773	
Qy	340	LysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSer	359	
Db	774	GCCCAATCATATCTCAACAA-----	794	
Qy	360	AspAlaLysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluI	377	
Db	795	GACTCGAAACGATGCTCAAGACGAGTCAAGAGAGTCCGATACGCTGGATTCATCAACAGC	854	
Qy	377	IleLeuAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAla	396	
Db	855	ACGTG---GCCAACGTCATCATCGACGGAAGTTCCCTAAGCATGCAATCCCATCTTCC	911	
Qy	397	AlaAspProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAsp	416	
Db	912	AAGAGTTCACGTCAGACGAAAGTCTCCGTATCATCAGACGATCTTACTCTTAACGCTCC	971	
Qy	417	Ile-----ProValSerAsnHisThrValGly	425	
Db	972	ATCGTGACAGTATCAGACACCGCATAGCCGCAACACCGGTTTCTCCAAATATTATCAAC	1031	
Qy	426	GluAsp-----GlyLeuLysSerSerLysAsnLysThrLysArg	438	
Db	1032	AAGCCTTTGAGGAAAAACCACTACGTCAGTGAAGAGAGTGAAGACAGACAGCAAAAAA	1091	
Qy	439	LysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLys	458	
Db	1092	GATCCACTCCAGCTGTTCCGCCAGTGACAC-----CAG	1127	
Qy	459	LysArgThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLys	478	
Db	1128	CCAACAATCGAGTGTGTAGTCCCAATTATGCGCAT-----AAG	1166	
Qy	479	LysValThr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsn	496	
Db	1167	AAGTTGACAAATGACCCCGTGATATCTGAAAAACACGACCTGAAAGAGTCCCAATCATG	1226	
Qy	497	GlyLeuAspThrAsnMetHisLysThrAspVal---CysGlnHisValSerGluIleSerT	516	
Db	1227	AGCATGACACG-----ACGAGCGTTCCACCGCTTCCACCTCTTAAATTCAGTT	1274	
Qy	516	hrGln-----ArgCysSerSerLysGlyL	524	
Db	1275	GTTCCACTTAAATGACTTCAATCCGACACCCACACGATACGATGTTCTCTTAAACAA	1334	
Qy	524	ysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysThrGlyGlyG	544	
Db	1335	GG-----AAAAATCACATCGCTGTCAAG---TCGTTTGGATATG	1371	
Qy	544	IuSerThrArgAsnGlyAlaAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetG	564	
Db	1372	AGCAGTCG-----TCOG	1383	
Qy	564	IuThrGluAsnSerValLeuSerHis---SerAlaLysValSerPro-----	578	

Db	1384	CGTCTGAAGACTCCATTGTGGCTCATGCTGGCTCAGGTGACTCCGCCGCAAAAACCTT	1443
Qy	579	--AlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLeuLysL	598
Db	1444	CTGTAATCATTCGCTG-----GAGAGAAGCATGGGAAGAATA	1482
Qy	598	ysylsLysGlnLysLeuGluValThr-----A	607
Db	1483	AGACATCAGAAATCCAGCGGTACACTCTGACCGCGGTGTGGCATGTGGCCGCAAAATGA	1542
Qy	607	rgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysA	627
Db	1543	GGGAGAGCTGAAAGATACGATGACTCGT-----CGAGCACAGA	1587
Qy	627	snGlnHis-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleA	643
Db	1588	ACGGCTATCTCGAACAATTCGAGACAGATTCTCTCTGTGCTCGGAATATCCGATAACA	1647
Qy	643	snArgIleGlnSerLysThrAlaasp-----AspAspCysValIleValAlaA	660
Db	1648	ACGAGCTCGACACATATCCACGCGAGATTGTGCGGAGTAGACATGGCAACATGCGCT	1707
Qy	660	lalysAspGlySerAspTyrAlaSerVal---PheAspThrAsnSerGlnGlnLys-	678
Db	1708	CCAAACAT---AGCGATATTCACCATTTGTTCGCCATCCACGTCTTCTTCCTCAAAGC	1764
Qy	579	-----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaL	695
Db	1765	CCGAGTCCCGAGTCGGTCTCCACATCAGTCGATCTCGATCTCGAGCGAA-----	1817
Qy	695	euThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrH	715
Db	1818	-----CAGGAGAAATGTATCAAACTTCGTCCAGTCCGACGCGCAACGACGTGGC-	1868
Qy	715	isleuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisH-	734
Db	1869	-----GCCCTGCCACCTCAACCTTCGGACACACATT	1899
Qy	735	-----AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspA	752
Db	1900	CGCTAAGATCCCGGATATCATCTATTCTCCA-----CACTTATCA-----GTGTCAG	1950
Qy	752	lalysLysLeuThrTrpGlnGlnPheLysAlaThrThrArg-----AsnSerProAlaA	770
Db	1951	CTGATAGGACACAATGTCTATCCACTCAGACTAGTCGACGACCTTCTTCACAAAAAC	2010
Qy	770	laThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisV	790
Db	2011	CAGCTATTCCGCCAATTT-----CATTCATTGAT-----CGTAATGCCACC	2055
Qy	790	alMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgT	810
Db	2056	TTCAAGAGTTTCACATCCACCGACACAGATGCGGCTCTCTTTGACCCCG-----	2105
Qy	810	yrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrM	830
Db	2106	-----AGACGGGTGCCGAATCGATCGATCGAAAT	2133
Qy	830	etGluAlaSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrP	848
Db	2134	ATGATTCTTCAGATCTCTACTCGCGCGGTCTCCGAGGTGGAAGCTCTACTGGTATCTATG	2193
Qy	848	rolysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaS	868
Db	2194	GAGAG-----ACGTTCCAATCCAGACTATCCGATGAAAAATCCCCCGCAC	2241
Qy	868	erpPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG	888
Db	2242	AT-----TCGCCCCAAAAGTGAGATGGGATCCCAACTA-----	2273
Qy	888	lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG	908

Db 2274 -----TCACTGGCTAGCAGCAGCATATGGATCT---CTCAATG 2310
Qy 908 lylsyleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgp 928
Db 2311 AGAAGTAC-----GAACATGCTATTCGGGACATGGCAGCTG 2346
Qy 928 rleu-----ArgProHisProArgValGlyV 937
Db 2347 ACTTGAGGTGTACAGAACACTGTGCTCACTAACCAAGAACACAGGAGAACTATGGAG 2406
Qy 937 alileGlySerLeuLeuGlnLysGluLeuAlaAsnTrpSerGluAsnCysGlyThrGlns 957
Db 2407 CATGTGTTGATCTTTTGGACAAAGCTTAGAAACACTCACTCAACACATGATGATCCA 2466
Qy 957 er-----GlyTyrLysLeuGlyVals 964
Db 2467 ACTTGAGGCTGAGAGGCAATACGATCAGCCAGCAGACATGCTCATTTGAGGGGATATTA 2526
Qy 964 erThrGlyLeuThrHisGlnMetAsnArgLysGlu----- 976
Db 2527 GCAATCATCTTGATCACTCAGCTCAGCTTAACGAGGGCTGTGAGCTTCCTGTC 2586
Qy 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerA 988
Db 2587 AACCATCTCTGGAATCAGTTGATCCCATGATCATGATGCTGCTGCGAAAGCA 2646
Qy 988 lalysTrpAsnAlaLeuGlnLysSerValSerSerSerAlaAspPheLeuSerAlaA 1008
Db 2647 GCAGCAGGAGAGATCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2687
Qy 1008 rgAsnSerileAlaGlnSerThrArgGlyLysGlyLysMetValHisProLeuAspA 1028
Db 2688 -----AAGAAGAGCTGATCGCTCTCA-----CTCTCCA 2718
Qy 1028 rgPheValArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrT 1046
Db 2719 AGTACCAGAG-----AAGAAGAGCTGATCGCTCTCA-----CTCTCCA 2766
Qy 1046 hrlsSerAsnAspAsnGluTyrMetAsp 1055
Db 2767 CAATTCGGATCTCAAGGAACCTTGAC 2795

RESULT 14
US-09-845-917A-53
; Sequence 53, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14.1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 4382
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-845-917A-53

Alignment Scores:
Pred. No.: 3,75e-09 Length: 4382
Score: 205.00 Matches: 232
Percent Similarity: 34.86% Conservative: 156
Best Local Similarity: 20.84% Mismatches: 379
Query Match: 3.71% Indels: 347
DB: 4 Gaps: 54

US-09-721-114-2 (1-1057) x US-09-845-917A-53 (1-4382)

Qy 110 LysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAla 129
Db 6 AAGAAATTTGGAGCACTACCCACATCCATATGACCCCGGGTTTCTAAATATCCCTTCG 65
Qy 130 LysGlnAsnGlyThrSerAspGlyCysSerileThrPheValArgSerThrPhe 147
Db 66 CCAGTGTGGCAGCTCAGCAACCGCTTCAGCAACTAACCCCAATTCACACTTCCACA 125
Qy 147 ----- 147
Db 126 ATGTCAACATCCAGCTTCAGACTCCAGTCCAGATATCGAAATATGATTCATCAAG 185
Qy 148 -----ValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSer 164
Db 186 ATTGTATCAAGCCCAAGACGCTGAGCTTAAACACCCCTCATCATCAACCACTTCATCA 245
Qy 165 GlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlnGlyAsnAsp 183
Db 246 AATATACAAATTCATTCGGT-----CCGTGAGCCGCTCGAGTGGCAATATAT 296
Qy 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAla 196
Db 297 GTGGCTCGAGATATTCACATCTCGAGAGAGCTTAGATATCATCATCAAGTACAGTCT 356
Qy 197 GluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspVal 216
Db 357 ATTTCGAATCTAAACCCGACCTACCTCCCAACTCCAAAACCTTCTAGACCAACACCCAG 416
Qy 217 AlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValPro 236
Db 417 CTAGTTCGTGTGCT-----ACAATACAAAATTCGGAGC----- 452
Qy 237 GlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeu 256
Db 453 -----TCAAGCTAGCGCTCGAAAGCGGTGAGCAGCCCAAACTT 494
Qy 257 SerGluValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGlu 273
Db 495 GCTTCTGTGAGACATATTGGAGCAAAACAGAGCCCGGATAACAGCGGTGGTGGTGGT 554
Qy 274 ThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnPro 288
Db 555 GGAATGCTGAATTAAGTATTTCAGTAGCAAAAACCCATCTTCCTCATCGAATAGCCCA 614
Qy 289 MetSerGlyLysGluArgAspGlnValAlaGlnGlnCysAsnLeuThrLysAspProLys 308
Db 615 CAACCTACGAGAAAGCGCGCGGTGCTCAACACAACTTTGTCGAAAATTCGCTGCC 674
Qy 309 ProVal-----SerGlyGlnLysCysGlnIleCysAsnGluProCysGluGluValVal 327
Db 675 CCAGTGAAGTGGCTGAG-----CCGCGACCACTAGCTG 713
Qy 328 LeuLysArgSerSerLysSerLys-----ArgLysThrAsp 339
Db 714 GGAAGTGCACGCTATGTCGAAGCTTTGTACGCCAAAAGTTTCTACCGTAAAACGGAC 773
Qy 340 LysLysLeuMetLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSer 359
Db 774 GCCCAATCATATCTCAACAA----- 794
Qy 360 AspAlaLysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluI 377
Db 795 GACTCGAAGCATGCTCAAGAGCAGTGAAGAGAGTCCGATACGCTGGATTCAACAGC 854
Qy 377 leileAsnAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAla 396
Db 855 AGCTC-----GCCAAGCTCATCATCGACGGAAGTCCCTTAAGCATGATTCACATCTCC 911
Qy 397 AlaAspProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAsp 416
Db 912 AAGAGTTCAAGCTCAGACGAGAAAGTTCGTCATCAGACGATCTTACTCTTAACGCTCC 971

QY 417 Ile-----ProValSerAsnHisThrValGly 425
Db 972 ATCTGACAGCTATCAGACAGCCGATAGCCGACACACCGGTTCTCCAAATATTATCAAC 1031
QY 426 GluAsp-----GlyLeuLysSerSerLysAsnLysThrLysArg 438
Db 1032 AAGCCTGTTGAGGAAACCAACACTCGCAGTGAAGAGGTGAAGACACAGCGAAGAAA 1091
QY 439 LysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLys 458
Db 1092 GATCCACTCCAGCTGTTCGCCACGTCACACC-----CAG 1127
QY 459 LysArgThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLys 478
Db 1128 CCAACAATCGAGTGTAGTCCAAATTATGCACAT-----AAG 1166
QY 479 LysValThr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsn 496
Db 1167 AAGTTGACAAATGACCCCGTGATATCTGAAAACAGCAACTGAAAAGCTCCCAATCAATG 1226
QY 497 GlyLeuAspThrAsnMetHisLysThrAspVal---CysGlnHisValSerGluLysSer 516
Db 1227 AGCATCGACACG-----ACGAGCTTCCACCGCTTCACCTCTAATAATCAGTT 1274
QY 516 hrGln-----ArgCysSerSerLysGlyL 524
Db 1275 GTTCCACTTAATGACTTCAATCGACACACCACCAACGATGTTCTTCTAAACAA 1334
QY 524 yThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyG 544
Db 1335 GG-----AAAATCACAATCGCTGCTCAAG--TCGTTTGGATATG 1371
QY 544 luSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetG 564
Db 1372 AGCAGTCG-----TCGG 1383
QY 564 luThrGluAsnSerValLeuSerHis---SerAlaLysValSerPro----- 578
Db 1384 CGTCTGAAGACTCCATGTGCTCATCGCTCGCTCAGTGACTCGCGCCGACAAAAACTT 1443
QY 579 --AlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysL 598
Db 1444 CTGTAATCATTCGCTG-----GAGAGAGGATGGGAAGAATA 1482
QY 598 yLysLysGlnLysLeuGluValThr-----A 607
Db 1483 AGACATCAGAAATCCAGCGGTACACCTCTCGACCGCGGTGTCGATGTCGCCAAATGA 1542
QY 607 xGlnLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysA 627
Db 1543 GCGAGAGCTGGAAGAATACGATGACATGACTCGT-----CGAGCACAGA 1587
QY 627 snGlnHis-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleA 643
Db 1588 ACGGCTATCTGACAACTTCGAAGACAGTTCTCTCTGCTGCTGGAATATCCGATAACA 1647
QY 643 snArgIleGlnSerLysThrAlaAsp-----AspAspCysValIleValAlaA 660
Db 1648 ACGAGCTCGACGATATCCACGAGCGAATTTGTCGAGTAGACATGCGCAACAGTCGCT 1707
QY 660 lalysAspGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys- 678
Db 1708 CCAAAACAT---ACGGACTATCCCACTTGTTCGCCATCCACAGTCTTCTTCCTCAAGC 1764
QY 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlnHisLeuAlaL 695
Db 1765 CCGAGTCCCGACGTCGCTCCACATCATGCTGATTCGATCTCGAGCAGAA----- 1817
QY 695 euThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrH 715
Db 1818 -----CAGGAAATGTGTAAACACTTCTGCCAGTCCCGAACAGCAGCGTGGC- 1868
QY 715 isLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis- 734

Db 1869 -----GCGCTGCCACCTCAACCTTCGGACAACATT 1899
QY 735 -----AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspA 752
Db 1900 CGCTAAGATCCCGGATACTCATCTCTATCTCCA-----CACTTATCA--GTGTCAG 1950
QY 752 lalysLysLeuThrTrpGluGlnPheLysAlaThrArg-----AsnSerProAlaA 770
Db 1951 CTGATAAGGACCAATGTCTATCTATCTCACTCACAGATAGTCAGACCTTCTTCAAAAAA 2010
QY 770 lathrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisV 790
Db 2011 CAACTTATTCAGCCAAATT-----CAITCACTTGT---CGTAATGTCACC 2055
QY 790 alMetGlySerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgT 810
Db 2056 TTCAAGAGTTTCATCTCCACCGACACAGATGCGGCTCTCTTGAGCCCG----- 2105
QY 810 yAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThM 830
Db 2106 -----AGCGGTGCGGAACCTCGATGTCGAAT 2133
QY 830 etGluAlaSerLysLeuCysAspArgAsnAlaGly-----GlnValValLeuTyrP 848
Db 2134 ATGATTTCTTCAGATCTCTACTCGGCGGTTCGAGGTGGAAGCTCTACTGCTATCTATG 2193
QY 848 roLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaS 868
Db 2194 GAGAG-----ACGTTCCAACTGACACAGACTATCCGATGAAAATCCCCGCGAC 2241
QY 868 exPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db 2242 AT-----TCTGCCAAAAGTGAGATGCGATCCCAACTA----- 2273
QY 888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db 2274 -----TCACTGGCTAGCAGCAGCATATGATCT---CTCAATG 2310
QY 908 lylLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisAsp 928
Db 2311 AGAAGTAC-----GAACATGCTATTTCGGACATCGGCACG 2346
QY 928 roLeu-----ArgProHisProArgValGlyV 937
Db 2347 ACTTCGAGGTGTACAGAACACTGTCGACTCACTAACCAAGAAACAGGAGAACTATGGAG 2406
QY 937 alLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnS 957
Db 2407 CATTTGTTGATCTTTTGAGCAAAAGCTTAGAAAACCTCACTCAACATGATGATCC 2466
QY 957 ex-----GlyTyrLysLeuGlyValS 964
Db 2467 ACTTGAAGCTCGAGAGGCAATACGATTCCAGCAGACATTCCTCATTTGAGGGATATTA 2526
QY 964 erThrGlyIleThrSerHisGlnMetAsnArgLysGlu----- 976
Db 2527 GCAATCATCTTCGATCCAACTCAGCTCAGTCTGCTAAAGAGGCGCTGGTAGCTTCTTCGTC 2586
QY 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerA 988
Db 2587 AACCATCTCTGGAAATCAGTTGCTATCCCATCGATCATCGATGTCGTCGTCGTAAGCA 2646
QY 988 lalysTrpAsnAlaLeuGlnLeuGlnLysSerValSerSerSerAlaAspPheLeuSerAlaA 1008
Db 2647 GCAAGCAGGAGAGATCAGCTTGAGCTCGTTGGCAAGAAC----- 2687
QY 1008 rgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspA 1028
Db 2688 -----AAGAGAGCTGGATCCGCTCTCA-----CTCTCCA 2718
QY 1028 rgPheValArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrT 1046

Db 2719 AGTTCAACCAAG-----AAGAACAAAGACTACGACGAAGACCATATGCCAT 2766

Qy 1046 hrllSerAsnAspAsnGluTyrMetAsp 1055
:|||||

Db 2767 CAATTTCGGATCTCAAGAACTCTTGAC 2795

RESULT 15

US-09-845-917A-50
; Sequence 50, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845.917A
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-845-917A-50

Alignment Scores:
Pred. No.: 4,15e-09 Length: 4642
Score: 205.00 Matches: 232
Percent Similarity: 34.86% Conservative: 156
Best Local Similarity: 20.84% Mismatches: 379
Query Match: 3.71% Indels: 347
DB: 4 Gaps: 54

US-09-721-114-2 (1-1057) x US-09-845-917A-50 (1-4642)

Qy 110 LysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAla 129
:|||||

Db 6 AAGAAATGGGCACTACCCACATCCATTATGACCACCCCGGTTCTTAATATACCCCTCG 65

Qy 130 LysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPhe 147
:|||||

Db 66 CCAGCTGCGCCAGTCAGCAACCGCTTCAGCAACTAACCCAAATCCCACTTCCACAA 125

Qy 147 ----- 147

Db 126 ATGTCAACATCCAGGCTTCAGACTCCACAGTCAGAAATATCGAAATTTGATTCATCAAG 185

Qy 148 -----ValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSer 164
:|||||

Db 186 ATGGTATCAAGCCAAAGACGCTGGACTTAAACCAACCTCATCATCAACCACTTCATCA 245

Qy 165 GlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp 183
:|||||

Db 246 AATAAACAATTCATTCCTG-----CCGTGAGCGGTTTCGAGTGGCAATAATAAT 296

Qy 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAla 196
:|||||

Db 297 GTTGCTCGAGTATCCACATCTCGAAGAGCTTAGAATCATCATCAAGTCACAGCTCT 356

Qy 197 GluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspVal 216
:|||||

Db 357 ATTTGGAATCTAAACCGACCTACCTCCCAATCCAAACCTCTTAGACCAACCAACCCAG 416

Qy 217 AlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValPro 236
:|||||

Db 417 CTAGTTGGTGTGCT-----ACAACTACAAAATCGAAGC----- 452

Qy 237 GlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeu 256
:|||||

Db 453 -----TCAAGAGTAGCCGCTCCGAAAGCCGTCGAGACCCCAAACTT 494

Qy 257 SerGluValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGlu 273
:|||||

Db 495 GCTTCTGTGAAGACTATTGGAGCAAAACAGAGCCCGGATAACAGCGGTGGTGGTGGT 554

Qy 274 ThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnPro 288
:|||||

Db 555 GGAATGCTGAATTAAGATTATTCAGTAGCAAAACCCATCTTCTTCATCGAATAGCCCA 614

Qy 289 MetSerGlyLysGluAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308
:|||||

Db 615 CAACTACGAGAAAGCGCGCGCTCTCAACAACAACTTTGCGAAAATCGCTGCC 674

Qy 309 ProVal---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluVal 327
:|||||

Db 675 CCAGTCAAAGTGGCTGAAG-----CCGCGGACCACTAAGCTG 713

Qy 328 LeuLysArgSerSerLysSerLys-----ArgLysThrAsp 339
:|||||

Db 714 GGAAGTGCCACGCTCTATGTCGAAGCTTTGTACGCCAAAGTTTCTTACCGTAAACGGAC 773

Qy 340 LysLysLeuMetLysGlnGlnHisSerLysArgThrAlaGlnAlaAspValSer 359
:|||||

Db 774 GCCCAATCATATCTCAACA----- 794

Qy 360 AspAlaLysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGlu 377
:|||||

Db 795 GACTCGAAACGATGCTCAAGAGCAGTGGAAGAGTCCGAGATACGCTCGATTCAACAGC 854

Qy 377 IleLeuAsnAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAla 396
:|||||

Db 855 AGGTC---GCCAAGCTCATCATCGACGGAGGTTCCTTAAGCATGATTCACATCTTCC 911
:|||||

Qy 397 AlaAspProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAsp 416
:|||||

Db 912 AAGAGTTCAACGTCAGACGAAAGTCTCCGTCATCAGACGATCTTACTCTTAACGCTCC 971

Qy 417 Ile-----ProValSerAsnHisThrValGly 425
:|||||

Db 972 ATCGTCAGAGCTATCAGACAGCGGATAGCGCAACACCGGTTTCTCCAAATATTATCAAC 1031

Qy 426 GluAsp-----GlyLeuLysSerSerLysAsnLysThrLysArg 438
:|||||

Db 1032 AAGCCTGTTGAGGAAAAAACCAACACTGCGAGTGAAGAGGATGAAAGACACACGCAAAAA 1091

Qy 439 LysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLys 458
:|||||

Db 1092 GATCCACCTCCAGCTGTTCCGCCACGTCACAC-----CAG 1127

Qy 459 LysArgThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLys 478
:|||||

Db 1128 CCAACAAATCGAGTGTGTAGTCCAATTATGGCACAT-----AAG 1166

Qy 479 LysValThr-----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsn 496
:|||||

Db 1167 AAGTTGACAAATAGACCCCGGTATATCTGAAACACAGAACCTGAAAGCTCAATCAATG 1226

Qy 497 GlyLeuAspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLysSer 516
:|||||

Db 1227 AGCATCGACAG-----ACGAGCGTTCACCGCTCCACCTCTAAATCAGTT 1274

Qy 516 hrGln-----ArgCysSerSerLysGlyL 524
:|||||

Db 1275 GTTCCACTTAAATGACTTCAATCCGACAAACCAACGACGATGCTTCTTCTAAACAA 1334

Qy 524 yThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGly 544
:|||||

Db 1335 GG-----AAAATCATCATCGCTGTCAAG--TCGTTTGGATAG 1371

Qy 544 LuSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetG 564
:|||||

Db 1372 AGCAGTCG-----TCG 1383


```
QY 564 luThrGluAsnSerValLeuSerHis---SerAlaLysValSerPro----- 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1384 CGTCTGAAGACTCCATTGGCTCATCGCTCGCTCAGTGACTCCGGCAGCAAAACTT 1443
QY 579 --AlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProIlyshL 598
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1444 CTGGTAATCATTCGCTG-----GAGAGAGGATGGGAAAGATA 1482
QY 598 ysLysLysGlnLysLeuGluValThr-----A 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1483 AGACATCAGATCCAGCGGCTACACTCTGAGCGCGGTGTGGGATGTGGCCAAATGA 1542
QY 607 rgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysA 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1543 GGGAGAAGCTGAAGATACAGTATGATCATCTCGT-----CGAGCACAGA 1587
QY 627 snGlnHis-----GluA'gGlnLeuMetThrGluThrAspCysSerAspIleA 643
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1588 ACGGTATCTCGAGCAACTTCGAGACAGTTCCTCTGCTCGGAATATCCGATACA 1647
QY 643 snArgIleGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaA 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1648 ACGACTCGAGACATATCCAGGAGCATTTGTCGGAGTAGACATGCAACAGTCGCT 1707
QY 660 laLysAspGlySerAspTyrAlaSerServal---PheAspThrAsnSerGlnGlnLys- 678
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1708 CCAAAACAT---AGCGACTATTCCTCCACTTTGTTCGCCATCCCAAGCTCTTCTCCTCAAAGC 1764
QY 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaL 695
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1765 CCGGAGTCCCGAGTCGCTCCATCATCAGTCGATTCGATCTCGACAGAA----- 1817
QY 695 euThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrH 715
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1818 -----CAGGAGAATGTGTACAAACTCTGTCCAGTGCAGCAACGCAAGCTGCG- 1868
QY 715 isLeuArgMetGluGluMetValThrIleAlaAlaSerProLeuPheSerHisHis- 734
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1869 -----GCGCTGCCACCTCAACCTTCGGACAACTT 1899
QY 735 -----AspAspGlnTyrIleAlaGluAlaProThrGluHisIrpGlyArgLysAspA 752
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1900 CGCTAAGATCCCGGATCACTCATCTATCTCCA-----CACTTATCA---GTGTCAG 1950
QY 752 laLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaA 770
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1951 CTGATAGGACACAAATGCTATGCACTCACAGCTAGTCGACGACCTTCTTCAAAAAAC 2010
QY 770 laThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisV 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2011 CAAGCTATTCAGGCCAATTT-----CATTCATTGAT---CGTAAATGCCACC 2055
QY 790 alMetGlySerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 810
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2056 TTCAAGAGTTCACATCCACCGGACGACAGAGTGGCGGCTCTCTGAGCCCG----- 2105
QY 810 yrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrM 830
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2106 -----AGCGGTGCCGAACTCGATGTGGAAT 2133
QY 830 etGluAlaSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrP 848
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2134 ATGATTCTTCAGGATCTACTCGGCGGCTTCGAGGTGGAGCTCTACTGCTATCTATG 2193
QY 848 roLysGluSerMetProAlaThrHisLeuLeuA'gMetMetAspProSerThrLeuAlaS 868
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2194 GAGAG-----ACGTTCCAACTGCACAGACTATCCGATGAGAAAAATCCCGGCAC 2241
QY 868 erPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2242 AT-----TCGCAAAAGTGAGATGGATCCCACTA----- 2273
QY 888 inTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
```

```
Db 2274 -----TCACTGGCTAGCAGACAGCATATGGATCT---CTCAATG 2310
QY 908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2311 AGAAGTAC-----GAACATGCTATTTCGGGACATGGCACGTG 2346
QY 928 roLeu-----ArgProHisProArgValGlyV 937
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2347 ACTTGGAGTGTACAGAACACTGTGACTCACTAACCAAGAACACAGAGAACTATGGAG 2406
QY 937 alLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnS 957
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2407 CATTTGTTGATCTTTTGGACAAAGCTTAGAAAACTCACTCAACACATTGATCGATCCA 2466
QY 957 er-----GlyTyrLysLeuGlyValS 964
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2467 ACTTGAAGCCTGAAGAGGCAATACGATTCAGGCGAGGACATTGCTCATTTGAGGGATATTA 2526
QY 964 erThrGlyIleThrSerHisGlnMetAsnArgLysGlu----- 976
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2527 GCAATCATCTTGCAATCCAACCTCAGCTCATGCTAACGAAGGCGCTGTGTGAGCTTCTGCTC 2586
QY 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerA 988
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2587 AACCATCTCTGGAATCAGTTGCATGCCATCGATCGATCGATCGCTGCTCGAAAAAGCA 2646
QY 988 laLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaA 1008
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2647 GCAAGCAGGAGAGATCAGCTTGAGCTCGTTGGCAAGAAC----- 2687
QY 1008 rgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspA 1028
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2688 -----AAGAAGAGCTGGATCCGCTCCTCA-----CTCTCCA 2718
QY 1028 rgPheValArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrT 1046
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2719 AGTTCACCAAGAG-----AAGACACAGAACTAGCACCAGCACATATGCCAT 2766
QY 1046 hrIleSerAsnAspAsnGluTyrMetAsp 1055
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2767 CAATTCGGATCTCAAGGAACCTCTTGAC 2795
```

Search completed: May 17, 2004, 10:07:18
Job time : 499 secs

Alignment Scores:

Pred. No.: 0 Length: 3896
 Score: 5226.00 Matches: 1057
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-721-114-2 (1-1057) x US-09-828-068-1 (1-3896)

QY 1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
 DB 241 ATGGAGATTGTTGCGAGTACATCAGGAGGAGCTCGTGTGTTGGACGCAATGTTATGCTT 300
 QY 21 AlaArgGlyThrGluValAlaProValLeuThrAlaThrProArgGln 40
 DB 301 GTCGTGTGGAACTGTGCTGTAGCCCGCAGTGTGGAGCTGACAGCGCGCTCGTCNG 360
 QY 41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
 DB 361 GATGCGAGCCGCTGAAAGCTGTGTAGACGAAACCGGCACAAACACCAATGCGAGCATTTCTCC 420
 QY 61 IleArgGlyThrValAlaLeuLeuGlnLysAspProLysPheCysSerLeuSerArg 80
 DB 421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAAGAGGATCCAAATTCGTCTCTATCTCG 480
 QY 81 IlePheHisAspGlnLysCysAspGluHisLysAlaSerSerProPheSerVal 100
 DB 481 ATTTTCCATGACCAGAAAATGTGATGAACACAAAGCTAGTTCGAAGCCCAATTTCTGTA 540
 QY 101 AlaLysPheArgThrAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
 DB 541 GCAAAATTTTCGACGATGGGATTTGCTCGAAGTGCCTTGGATTAAGTTGAAAACCTTCAGATAAT 600
 QY 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
 DB 601 GGNACAGCCACCAAGAACTCTCCCGCAGACAGAGATGGCACAAAGTGTGTTGCTCCATC 660
 QY 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
 DB 661 ACATTTGTTTCGGAGCATTGTCGCTGCTAGTGTGTTGGTTCCCAAAAAGTGTCTCTTAGC 720
 QY 161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
 DB 721 ACACATCATCTCAAGGGAAGAAATGCTGATGATCACTCTTCCAAAGAGTGTGCAGAA 780
 QY 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaSerThr 200
 DB 781 GGCATATGACTCCAAATGCAATCGCCTCTGGCAAGAAATGGAGCTGCTGAGGCCAATACT 840
 QY 201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
 DB 841 GATTCACCAATGAAGATTTCGACGGCCAGCCCAAAATATGATGTGCGACCAATGTC 900
 QY 221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
 DB 901 TCTGAGGACACACTTCTGTTGATGTGGGCTTTTACCCTGAAGTTCCCCAGATTACATG 960
 QY 241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
 DB 961 CACATAGAGTAATGTTGGTGCATCAACCTCCATCCCTCCAAACTTCTTGAAGTGGTC 1020
 QY 261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
 DB 1021 CTCAAAAGAAATGAAGATGAATGGAAAACCTGAAGAGACTCTTGTGCTGAGCAGTGC 1080
 QY 281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
 DB 1081 AATTTGACCAAGATCTCAACCAATGTCTGGAAAGGAAACGTGATCAGCTTCTGCTGAGCAG 1140
 QY 301 CysAsnLeuThrLysAspProLysProValSerGlyLysGlnLysCysGluGlnIleCysAsn 320

DB 1141 TGCAATTTGACCAAAAGATCCGAAACCAAGTGTCTGGCGAGAAATGTCGACAGATCTGCAAT 1200
 QY 321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
 DB 1201 GAGCATGTGAAGAAGTTGTTCTCAAGAAGAGCTCCAAATCTAAGAGAGACCGGATGAAG 1260
 QY 341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
 DB 1261 AAGTTGATGAAGAAGCAGCAGCAGCAAGAAACGCACTGCCAGGCTGATGTTTCAGAT 1320
 QY 361 AlalysLeuCysArgArgLysProLysValArgLeuLeuSerGluLulleIleAsnAla 380
 DB 1321 GCAAGCTTTGTCGAGAAAGCCAAAAGAGTGGCGCTTCTATCAGAAATATATAATGCT 1380
 QY 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
 DB 1381 AACCAAGTTGAGGATTTCTAGAAGTGACGAGTTTCATCGTGAATGCGCTGATCCCTGT 1440
 QY 401 GluAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
 DB 1441 GAGATGATGAAGTACATCCCGTCCCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1500
 QY 421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTrp 440
 DB 1501 AACCATACAGTGGAGAGATGGGTAAATCAAGTAAGACACAGACAAACGCAATAC 1560
 QY 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
 DB 1561 TCTGATGTTGTAGATGATGATCATCTTATGAATCGGTGATGATGATGATGATGATGATG 1620
 QY 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyValAsnLeuSerAsnLysVal 480
 DB 1621 ACTGGAGTGTGATCATCACAGTTGCTTCATCCAGCTGGAAATTTGACACACAAAGATG 1680
 QY 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
 DB 1681 ACACCACTGCGAGTACTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 QY 501 AsnMetHisLysThrAspValCysGlnHisValSerGluLysSerThrGlnArgCysSer 520
 DB 1741 AATATGATGAACAGATGTTGTGACATGTTATCAGAAATCTCCACACAGAGTGTCTCA 1800
 QY 521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
 DB 1801 TCAAAGGGGAAACAGCGGTTTGAGTAAGGGGAAACACATTCAGCTGCTAGTACCAA 1860
 QY 541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
 DB 1861 TATGCTGTGAAGACACCCAGAAATGCGTCAGAACATACATGTTACTCAGCGCAGAGATCAA 1920
 QY 561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
 DB 1921 TGCCAGATGGAAACCGAAACTCTGTTCTGAGTCACTCCGCAAAAGTTTCTCCAGCTGAG 1980
 QY 581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLys 600
 DB 1981 CATGATATCCAAATATATGCTGACCTTCATGAGCAGAGTCTACCCAAAGAAAGAAAG 2040
 QY 601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
 DB 2041 CAATAACTTGAGTCACTCTGTGAATAACACACCATGATGATGATGATGATGATGATG 2100
 QY 621 ValGluLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
 DB 2101 GTTCAACTGCTAGTAAACCCAGCATGAGAGCGAGCTTATGATGATGATGATGATGATGATG 2160
 QY 641 AspIleAsnArgIleGlnSerLysThrAlaAspAspCysValIleValAlaAla 660
 DB 2161 GACATCAACGCTATTCAATCCAGACAACTGCTGATGATGATGATGATGATGATGATGATG 2220
 QY 661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
 DB 2221 AAGGATGTTTCAGATTATGATCAAGTGTCTTTCAGACACTAATTTCCCAACAGAGTCTTGT 2280

QY 681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
DB GCATCCCAAGTACACAGAAAGAGTACAGGGTCTTTGGCATTTGACCAACAGAGTCT 2340
QY 701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrHisLeuArgMetGluGlu 720
DB CCATCCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGGGATGGAAGA 2400
QY 721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspGlnThrIleAla 740
DB ATGGTCACATATTGCTGCAAGCTCACCACATTTTTCATCATCATCATCATCATCATAT 2460
QY 741 GluAlaProThrGlnHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
DB GAAGCACCACCTGAACATTTGGGCGGTAGGAGCGCAAGAGCTAAGCTGGGAGCAATTT 2520
QY 761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
DB AAGGCCACTACAGAAATTTCTCCAGCAGCAACATGTGGTCTCAATTTAGACCTGGTATC 2580
QY 781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTrpAlaSerArg 800
DB CAGCAGTTCAGTTCAGTCTACTCATGTCATGCGATCTCCAGCAATTTATGCATCTCC 2640
QY 801 GlnProValIleAlaProLeuAspArgTrpAlaGluArgAlaValAsnGlnValHisAla 820
DB CAACCACTAATTCGGCCACTGGACCGCTATGCTGAAGAGCGGTAAACCAAGTCCATGCA 2700
QY 821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
DB AGAAATTTCCAGCAACATAGCAACCATGGAAGCGAGTAAGTTATGTATCGAGAAAT 2760
QY 841 AlaGlyGlnValValLeuTrpProLysGlnSerMetProAlaThrHisLeuLeuArgMet 860
DB GCTGGCAAGTACTCTGTATCTCTAAGAAATCATGCTCGCAGCATCTCTTGAGAATG 2820
QY 861 MetAspProSerThrLeuAlaSerPheProAsnTrpGlyThrSerSerArgAsnGlnMet 880
DB ATGGATCCATCAACATTAGCACTTCCCACTATGGAACCTCTAGCAGGACCAAGATG 2880
QY 881 GluSerGlnLeuHisAsnSerGlnTrpAlaHisAsnGlnTrpLysGlySerThrSerThr 900
DB GAGCTCTCACTTCATATTTCTCAGTATGCACATATCAGTACAAAGATCAACCAACACA 2940
QY 901 SerTrpGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
DB TCATATGGCAGTAACCTGAATGGAAGATTTCCATTGACATTTGGAAGACTTATCACCGCAT 3000
QY 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
DB CAGCTGCATGATCTGCACAGACCTTTACGCCACATCTTAGAGTTGGTGTGCTGGCTCC 3060
QY 941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTrpLys 960
DB TTGCTCGAAGAGAAATTTGCAAACTGGTCGGAGAACTGTGGCACAATCTGTTTATAAG 3120
QY 961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
DB TTAGGAGTGTCAACAGAAATATCATCGCATCAGTGAACAGAAAGAAATTTTGAAGCC 3180
QY 981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
DB CTGAATTTCTGGAATGTTTTCAGCAAAATGAAATGCAATGCAATTCAGTTGGGTTCTGTAGCTCC 3240
QY 1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
DB AGTGCAGATTTTATATCAGCGAGAAACAGCATAGCTCAATCTTGGACCAAGAGCGAGGT 3300
QY 1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
DB AAATGGTTCATCCCTTGGATCGTTTGTGAGACAGATATCTGTATATCACTACAGAAC 3360

QY 1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTrpMetAspTrpArg 1057
DB CCAGCTGATTTTACTACATCAATCAGTACGATACAGATATATGATACCGC 3411
RESULT 2
US-10-447-135-1
; Sequence 1, Application US/10447135
; Publication No. US20030199684A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Hirohiko
; APPLICANT: Hirochika, Hirohiko
; APPLICANT: Miyao, Akio
; TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
; FILE REFERENCE: MAPF-1 DIV
; CURRENT APPLICATION NUMBER: US/10/447,135
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US 09/721,114
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JAPAN 2000-149106
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4310
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (655)..(3828)
US-10-447-135-1
Alignment Scores:
Pred. No.: 0 Length: 4310
Score: 5526.00 Matches: 1057
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-721-114-2 (1-1057) x US-10-447-135-1 (1-4310)
QY 1 MetGluIleValAlaValAspGlnGluGlyAlaArgValAlaGlyThrAsnCysMetLeu 20
DB 655 ATGGAGATTTGTTGCACTAGATAGTACAGAGGAGCTGCTGTTGTTGGCAACATGTATGCTT 714
QY 21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
DB 715 GCTCTGTGTGGAACCTGGTGTGTAGCGCCAGTGTGGAGCTGACAGCGCTCGTCAG 774
QY 41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
DB 775 GATGACGCGCTGAGCTGTGTAGACGACCGGACCAACCAATGCGAGCATTTCTCC 834
QY 61 IleArgGlyTrpValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
DB 835 ATAGAGGTATGTGTGCTCTTCTTCCAGAAAGAGGATCCAAAATTTCTGCTCTATCTCGG 894
QY 81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
DB 895 ATTTTCCATGACCAAGAAATGTGTATGACACAAAGTAGTTTCAGGCCCATTTTCTGTA 954
QY 101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
DB 955 GCAGAAATTTCCAGCATGCGATTTCTTCCAGATGCTTTGGATAAGTTGAAACCTCAGATAAT 1014
QY 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
DB 1015 GGAACAGCAGCAAGAACTCTTCCCCCAAGCAAGATGGCAAGATGATGTTGCTCCATC 1074
QY 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
DB 1075 ACATTTGTTCCGAGCACATTTTGTGCTGTAGTTGTGTTGCTTCCAAAAAGTGTCTCTAGC 1134

161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
1135 ACACAATCATCTCAAGGGAGAAATGCTGATAGTCAACTCTTCCAAAGAGTGTGCAGAA 1194
181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyValAlaAlaAsnThr 200
1195 GGCATGACTCCAAATGCAATCGGCTCTGCGAAGATGGAGCTGCTGAGGCCAATACT 1254
201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnThrAspValAlaAlaVal 220
1255 GATTCCACCAATGAAGAATTGCAAGGGCCAGCCCAAAATATATGATGTGCGACAAATGTC 1314
221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
1315 TGTGAGGACCAACACTTCTGTGATGTGGGCTTTTACCTGAAGTTCCTCCAGATTACTAGG 1374
241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCCAAACTTCTGAACTGGTC 1434
261 LeuLysArgAsnGluAspGluAsnGlySerThrGluGluThrLeuValAlaGluGlnCys 280
1435 CTCAAAGAAATGAAGATGAAATGGAAAACTGAAGACACTCTGTGCTGAGCAGTGC 1494
281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
1495 AATTTGACCAAGATCTCTAACCAATGTCTGGAAGGAACGTGATCAGGTTGCTGAGCAG 1554
301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
1555 TGCAAATTTGACCACCAAGATCCGAACAGTGTCTGGCCAGAAATGTGAGCAGATCTGCAAT 1614
321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
1615 GACCCATGTGAAGAGTTGTTCTCAAGAGAGCTCCAAATCTAAGAGGAAGCAGATAG 1674
341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
1675 AAGTTGATGAAGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1734
361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
1735 GCAAGCTTTGTCCGGAAGGCCAAAGAGTGGCTGCTCTATCAGAAATTTAATAGT 1794
381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
1795 AACCAAGTTGAGGATCTAGAGTGCAGAGTTCATCGTCAAAATGCCGCTGATCCCTGT 1854
401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
1855 GAGGATGATAGAAGTACCATCCCGGTCCTCCGATGGAAGTAAAGCATGATATCTGTAGC 1914
421 AsnHisThrValGlyLysAspGlyLeuLysSerSerLysAsnLysThrLysArgLysThr 440
1915 AACCATACAGTGGGAGAGATGGTTAAATCAAGTAAGAACCAAGACAAACGCAAAATAC 1974
441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
1975 TCTGATGTTGTAGATGATGATCATCTTATGAATGGCTGAATGGAAAAAGAAAGA 2034
461 ThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
2035 ACTGGAGTGTGCATCACAGTTGCTCATCCAGCTGGGAATTTGACCAACAAAAAGTG 2094
481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
2095 ACACCCACTGGAGTACTCAGCATGATGATGAGATGATGATGATGATGATGATGATGATG 2154
501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
2155 AATATGATATACAGAGTGTCTGTCAGCATGTATCAGAAATCTCCACAGAGGTGTCTCA 2214
521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540

2215 TCAAAGGGGAACACAGCGGTTTGGTAAAGGGGAACACATTCAGCTGCTAGTACCAA 2274
541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
2275 TATCGTGTCAAGACACAGAAATGGTCAGACATACATGCTACTCAGCGCAGAGATCAA 2334
561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
2335 TCCAGATGGAAACCGAAACTCTGTTCTGAGTCACTCGGCAAAAGTTTCTCCAGCTCAG 2394
581 HisAspIleGlnIleMetSerAspLeuHisGlnSerLeuProLysLysLysLysLys 600
2395 CATGATATCCAAATATATGCTGACCTTTCATGAGCAGAGTCTTACCCAAAGAAAGAAAG 2454
601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
2455 CAAAACCTTGAAGTGACTCTGTGAANAACGACCATGATAGATGATATCCCATGGAATTT 2514
621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
2515 GTTGAACCTGTAGCTAAANACCAGCATGAGAGGCGAGCTTATGACTGAGACTGATTTGTCT 2574
641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspCysValIleValAlaAla 660
2575 GACATCAACCGTATTCATCCACACACTCTCTGATGATGATGATGATGATGATGATGATG 2634
661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
2635 AAGGATGGTTCAGATTATGTCATCAAGTGTGTTGACACTAATTCCTCCACAGAGTCCTTG 2694
681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisIleAlaLeuThrThrGlnGluSer 700
2695 GCATCCCAAGATACACAGAGAGTTACAGGCTCATTTGGCATTGATGACCAACAGAGTCT 2754
701 ProHisProGlnAsnPheGlnSerThrGlnGlnGlnThrHisLeuArgMetGluGlu 720
2755 CCACATCTCTCAGACTTTCAGTCTACTCAGGAACAGCAGACACATTTCCGCGATGAGAA 2814
721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
2815 ATGGTCACTATTGCTGCAAGCTCACCACATATTTTCCATCATCATGATGATGATGATG 2874
741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
2875 GAGCCACCACTGAACAATTTGGGCGGTAAAGACGCAAGAGCTAACCTGGGAGCAATTT 2934
761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
2935 AAGGCCACTTACAAAGAAATTTCCAGCAGCAACATGTGTGTCTCAATTTTAGACCTGATC 2994
781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
2995 CAAGCAGTGTGACTTGACTTCTACTCATGTCTGGGATCTTCCAGCATTTATGCACTCGC 3054
801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
3055 CAACCCAGTAAATTCGCCACTGCGACCGCTATGCTGAAGAGCGGTTAACCGATCCATGCA 3114
821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgAsn 840
3115 AGAAATTTTCCAAAGCACAAATAGCAACCCATGGAGCGAGTAAGTTATGTGATCGGAGAAAT 3174
841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
3175 GTTGGACAGTAGTCTTGTATCTTAAGAATCCATGCTCGCCTGCGAGCATCTTCTGAGAATG 3234
861 MetAspProSerThrIleAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
3235 ATGGATCCATCAACATTAGCAAGCTTCCCAACTATGGAACTTCTAGCAGGAACCAAGATG 3294
881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900

Db 3295 GAGTCTCAACTTCATATTTCTAGTATGCACATAATCAGTACAAAGAGTCAACCCAGCACA 3354
Qy 901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
Db 3355 TCATATGCGAGTAACTGAATGGAAGATTCATTCATTCGAGACTTATCAGCGCAT 3414
Qy 921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
Db 3415 CAGCTGCATGATCTGCACAGACCTTTACGCCACATCTAGTGTGCTGGCTCC 3474
Qy 941 LeuLeuGlnLysGluIleAlaAsnTyrSerGluAsnCysGlyThrGlnSerGlyLys 960
Db 3475 TTGCTGCAGAGGAAATTTGCAACTGTGCGAGACTGTGCGCACCAATCTGGTTATPAG 3534
Qy 961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
Db 3535 TTAGAGTGTCAACAGCAATAACATCGCATCAGATGAACAGAAAGAACATTTTGAAGCC 3594
Qy 981 LeuAsnSerGlyMetPheSerAlaLysTyrAsnAlaLeuGlnLeuGlySerValSer 1000
Db 3595 CTGATTTCTGGAATGTTTTTCAAGAAATGGAATGCATTCAGTGGGTTCTGTAGCTCC 3654
Qy 1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerThrArgGlyLysGly 1020
Db 3655 AGTGCAGATTTTATCAGCGAGGAACAGCATAGCTCAATCTGGACCAAGGAGGT 3714
Qy 1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
Db 3715 AAAATGTTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATTAACCTAACAGAAC 3774
Qy 1041 ProAlaAspPheThrThrIleSerAsnAspAsnGlnLysMetAspTyrArg 1057
Db 3775 CCAGCTGATTTTACTACAACTAGTAACGTAACGATATATGATTAACCGC 3825

RESULT 3
US-10-447-135-3
; Sequence 3, Application US/10447135
; Publication No. US200301996841
; GENERAL INFORMATION:
; APPLICANT: Hirochika, Hirohiko
; APPLICANT: Yamazaki, Muneko
; TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
; FILE REFERENCE: MAFF-1 DIV
; CURRENT APPLICATION NUMBER: US/10/447,135
; PRIOR APPLICATION NUMBER: 2003-05-27
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JAPAN 2000-149106
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 9455
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-447-135-3

Alignment Scores:
Pred. No.: 0 Length: 9455
Score: 5191.00 Matches: 1055
Percent Similarity: 68.06% Conservatives: 0
Best Local Similarity: 68.06% Mismatches: 2
Query Match: 93.94% Indels: 495
DB: 15 Gaps: 3

US-09-721-114-2 (1-1057) x US-10-447-135-3 (1-9455)
Qy 1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
Db 4223 ATGAGATTTGTTCAGTAGATCAGGAGGAGCTCGTGTCTTGGGACGAAGTGTATGCTT 4282
Qy 21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40

4283 GCTCGTGGTGGAACTGGTGTCTAGCGCCAGTGTGGAGCTGACAGCGCGCTCGTCAG 4342
Qy 41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
Db 4343 GATGCGAGCCGCTGAAGCTGTGTAGACGAACCCGCGACAAACACCAATGCGAGCACTTCTCC 4402
Qy 61 Ile----- 61
Db 4403 AT- AAGGTAATCATTTCTCTGTATTTTCCAATTCAGTATCGCGTTGTGGATGAATAATGA 4461
Qy 61----- 61
Db 4462 TCGGCATGTCAAGCCCATATTCGACTGTGTTGATGGAAGAGATATGATTGATCGTGGTTTTG 4521
Qy 61----- 61
Db 4522 CACAGTTTGTCTGTGGGACTTATATGGTCACTGTGTTTGTGACGATCGTATACACTGGGTC 4581
Qy 61----- 61
Db 4582 GACATGCTTATGCACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 4641
Qy 61----- 61
Db 4642 CTAGCCATGTGAAGTCAATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAGCAAAATATA 4701
Qy 61----- 61
Db 4702 TATAATATTTGCATATATGTTGGTGTCTCATGCTATCTTTATCTCTACGTACATCCATTA 4761
Qy 61----- 61
Db 4762 ATATCTTCAATGTATGAATCTGAGCACATGATGTGAGTGTCTACACATATGCGATGCTGT 4821
Qy 61----- 61
Db 4822 ATGTGTGTTTCAATAGGTGTTTCATCATATTTGTTGTTGGGGTGGCGATGCAATTTAT 4881
Qy 61----- 61
Db 4882 CAGGCCATGCTGTAGGCTGTAGATATTTGTGTTGTATATTTCTCTGTGAGACAA 4941
Qy 61----- 61
Db 4942 GCTGATTAATAAGAAATTAACCTTTTGGGGTACACTCATATATTTGGGCCCTACATTTT 5001
Qy 61----- 61
Db 5002 TGTAAATCAATTTTCCCTTTGTGCTGAGGTTTCAGCATAAACCTTTTATCATAAAGCATGTT 5061
Qy 61----- 61
Db 5062 TACATCCTTAGGAGATTTCTAGAACTGATGTTTCTTCATATTTGCAATATGTTGATTG 5121
Qy 61----- 61
Db 5122 ATAGTCATTAATTTTAAAGCCTTTTCAATTTGTTAGAGATTTAGAGATGATATATAT 5181
Qy 61----- 61
Db 5182 CAACCATAGACTGTGCACGTTTTGTTTAACTTTCTAGAACTAATTAGATTAATTTT 5241
Qy 61----- 61
Db 5242 TTGTAGTTTATCCTGTGTATGTTTGTATTTATCTTTTGAATTTCAAACCTGCAATACTAGA 5301
Qy 61----- 61
Db 5302 TTATCTTGAAGTCTCTCTTTTCTGTGAGCTGTCAAGCTATGATGAATGACCTACCTCCC 5361
Qy 61----- 61

Db 5362 AGCATCTTTAGATTATGTAGGGCCCTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 5421
QY 61 ----- 61
Db 5422 ACGCAATGCTATATATATGTGCCATGTCATCTTTAATGATAATCTTTATTTCTTGT 5481
QY 62 ----ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
Db 5482 ACCAGAGGGTATGTTGCTCTCTTCAGAAAGAGATCCAAAATTTCTGCTCTATCTCGG 5541
QY 81 ILlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerProPheSerVal 100
Db 5542 ATTTTCCATGACCAAGAAAATGATGAACACAAAGCTAGTTCAGGCCCATTTCTGTGA 5601
QY 101 AlaLysPheArgArgTirpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
Db 5602 GCAAGATTTTCGACGATGGGATGCTCGAAGTGCTTGGATAAGTTGAAACACTTCAGATAAT 5661
QY 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db 5662 GGNACAGCACCAAGAACTCTTCCCGCAAGACAGAAATGGCACAAAGTGATGGTTGCTCCATC 5721
QY 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
Db 5722 ACATTTGTTCCGAGCACTTTTGGCCCTGCTAGTGTGGTTCCCAAAAAGTGTCTCTAGC 5781
QY 161 ThrGlnSerSerGlnGlyLysAlaLeuAspArgSerThrLeuProLysSerValGlnGlu 180
Db 5782 ACACAATCATCTCAAGGGAAGATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 5841
QY 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyValAlaLeuAlaAsnThr 200
Db 5842 GGCATAGACTCCAAATGCCAATGCCCTTCTGGCAAGATGGAGTCTCTGAGGCCAATACT 5901
QY 201 AspSerProMetLys----- 205
Db 5902 GATTCCCAATGAA - AGGTATGATAGATGTAGAGCCCTTCAAATTCCTAAGTAGGATTTT 5960
QY 206 ----- AspLeuGlnGlyProAla 211
Db 5961 ATTTAGGTATAGAAATAAATAATGTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 6020
QY 212 GlnAsnTyrAspValAlaAlaLeuValSerGluAspAsnThrSerValAspValGlyAla 231
Db 6021 CAAATTAATGATGGCAGCAAAATGTCTCTGAGGACAACTCTCTGATGTGGGGCT 6080
QY 232 LeuProGluValProGlnIleThrTirpHisIleGluValAsnGlyAlaAspGlnProPro 251
Db 6081 TTACCTGAAGTTCCCGAGATTACATGCGACATAGAAATGAATGATGTCAGATCAACTCCA 6140
QY 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGlnAsnGlyLysThr 271
Db 6141 TCCACTCCAAAACCTTCTGAAGTGGTCTCTCAAAGAAATGAAGATGAAAATGGAAAAACT 6200
QY 272 GluGluThrLeuValAlaGlnCysAsnIleuThrLysAspProAsnProMetSerGly 291
Db 6201 GAAGAGACTCTTGTCTGAGCAGATCTGCAATTTGACCAAGAGTCTTAACCCAAATGCTGGA 6260
QY 292 LysGluArgAspGlnValAlaGlnCysAsnLeuThrLysAspProLysProValSer 311
Db 6261 AAGGAAGTGATCAGGTTGCTGAGCAGTGCATTTTCACCAAGATCCGAAACCAAGTGTCT 6320
QY 312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db 6321 GGCCAGAAATGTGAGCAGATCTGCAATGAGCCATGATGAAGAGTGTGTTCTCAAAAGAAAGC 6380
QY 332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
Db 6381 TCCAAATCTAAGGAGGAGACGGATAGAAGTTGATGAAGACGAGCAGCAGCAGCAAGAA 6440
QY 352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
Db 6441 CGCACTGCCCGGCTGATGTTTTCAGATGCAAAAGCTTTTGTGGGAGAAAGCCAAAAGAGTG 6500

QY 372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db 6501 CGGCTTCTATCAGAAATATATAATGCTTAACCCAGGTTTCCAGAGTTCTAGAGTTCAGCAAGTT 6560
QY 392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db 6561 CATCGTGAATAATGCGCTGATCCCTGTGAGGATGATAGAAGTACATCCCGGTCCCGATG 6620
QY 412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db 6621 GAAGTAAGCATGATATTCCTGTAGCAACCATACAGTGGGAGAGATGGGTTAAAAATCA 6680
QY 432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db 6681 AGTAAGAACCAAGACAAACGCAATATCTCTGATGTTGTAGATGATGATCATCATCTTATG 6740
QY 452 AsnTirpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db 6741 AACTGGCTGAATGGAAAAAGAAAGAACTGGAAGTGTGCATCACAGTTGTCTCATCCA 6800
QY 472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db 6801 GCTGGAAATTTGGCAACAAAAAGTGCACCCACTGCGAGTACTCAGCATGTGATGAG 6860
QY 492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511
Db 6861 AATGATCTGAAATGGTCTTGCACACAAATATGCATTAAGACAGATGCTCTGTGAGCATGTA 6920
QY 512 SerGluIleSerThrGlnArgCysSerSerLysLysLysThrAlaGlyLeuSerLysGly 531
Db 6921 TCAGAAATCTCCACACAGAGTGTCTCAAAAGGGGAAACAGCGGGTTTTCAGTAAGGGG 6980
QY 532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db 6981 AANAACACATTCACTGCTGTAGTACCAATATGCTGTGTGAACACCAAGAAATGGTCAGAAC 7040
QY 552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db 7041 ATACATGTACTCAGCGCAGAAAGATCAATGCCAGATGGAAACCGAAACCTCTGTTCTGAGT 7100
QY 572 HisSerAlaLys----- 575
Db 7101 CACTCGGCAAGAGTACGNAATTTTGTGAATCATGAGGAATTTTGTCTTTTAAATGACTG 7160
QY 575 ----- 575
Db 7161 AATCAACATTTATCTGTATGAAGCAATAATATGCTGCATAAACAATGTTAAGAAATATGC 7220
QY 575 ----- 575
Db 7221 ATACAAATGTTTATTATATGAGCTTTCCACTGTCTCTTCTTACTTATGTTTGTACTCTTT 7280
QY 575 ----- 575
Db 7281 TTGTGTGCGGTGCATGTGTCATG 7340
QY 575 ----- 575
Db 7341 ACGTGGTGGCGCAATATCTTTTATAGACTCAATATATAGTGAATGTAATGGACTGACA 7400
QY 576 -----ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db 7401 TTTTCTCTATTTCTCATCTCAGGTTTCTCCAGCTGAGCATCATATCCAAATATATGCTGA 7460
QY 588 pLeuHisGluGlnSerLeuProLysLysLysGlnLysLeuGluValThrArgGln 608
Db 7461 CTTTCATGAGCAGAGTCTTACCCAGAGAGAAAGAAAGCAAAACTTGAAGTACTGCTGGA 7520
QY 608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAspGln 628
Db 7521 AAACGACCCATGATAGATGACATCCCCCTGGAATATGTTGGAATGCTAGCTTAAACCA 7580

QY 628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerIly 648
Db 7581 GCATGAGAGCGAGCTTATGACTGAGACTGATTGCTGACATCAACCGTATTCAATCCAA 7640
QY 648 sThrThrAlaAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlase 668
Db 7641 GACAACTGCTGATGATGATGTTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCAATC 7700
QY 668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGln 688
Db 7701 AACTGTGTTTGACATTAATCCCAACAGAGTCTCTTGSCATCCCAAGTACACAGAAGA 7760
QY 688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGlnSerProHisProGlnAsnPheGlnSe 708
Db 7761 GTTACAGGGTCAATTTGGCAATTCACACAGAGTCTCCACATCTCCAGAACTTTTCAGTC 7820
QY 708 rThrGlnGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
Db 7821 TACTCAGGAACACAGACACATTTGGGATGGAGAAATGGTCACTATTGCTGCAAGCTC 7880
QY 728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
Db 7881 ACCACTATTTCACATCATGATGATGATATATGCTGAAGCACCACTCAACATTTGGG 7940
QY 748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
Db 7941 CCGTAAGGACGCAAGAGCTAACGTGGGAGCAATTTAAGGCCACTACAGAAATTTCTCC 8000
QY 768 oAlaAlaThrCysGlyValaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
Db 8001 AGCAGCAACATGTTGGTGTCTCAATTTAGACTGGTATCCCAAGCACTTGACTTCACTTAC 8060
QY 788 rHisValMetGlySerSerSerSerThrThrAlaSerArgGlnProValIleAlaProLeuAs 808
Db 8061 TCATGTCATGGGATCTTCAGCAATTTATCATCTCCCAACCAAGTAATTTGCGCCACTGGA 8120
QY 808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
Db 8121 CCCTATGCTGAAGAGCGGTAAACAGGTCCATGCAAGAAATTTTCCAGCAACATATAGC 8180
QY 828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
Db 8181 AACCATGGAAGCGAGTAAGTATGTGATCGAGAAATGCTGGCAAGTAGTCTGTGATCC 8240
QY 848 oLysGlnSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrIleuAlase 868
Db 8241 TAAAGAAATCCCTGCGGACGATCTTCGAGATGATGATGATCAATCAACATAGCAAG 8300
QY 868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
Db 8301 CTTCCCACTATGGAATCTTAGCAGGAACCAAGATGGAGTCTCAACTTCATATTTCTCA 8360
QY 888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
Db 8361 GTATGCACATTAATCAGTACAAAGGATCAACACACATCATATATGGCAGTAACCTGAATGG 8420
QY 908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
Db 8421 AARGATTCATTAATTCAGACACTTATCAGCGCATCAGTGCATGATCTGCAGAGACC 8480
QY 928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
Db 8481 TTTACGCCCACTATCAGATGGTGGTGGTCTCTCTGCTGCAGAAAGGAATTCGAAA 8540
QY 948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
Db 8541 CTGGTCCGAAACCTGTCGACACAACTGTTTATAGTTAGGAGTGTCCACAGAAATAC 8600
QY 968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
Db 8601 ATCCCATCAGATGAACAGAAAGAACATTTTGAAGCCCTCAATCTCGAATGTTTTCAGC 8860
QY 988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerValSerSerAlaAspPheLeuSerAlaAr 1008

Db 8661 AAAATGGAAATGCATTCAGTTGGGTTCTGTAGCTCCAGTCAGATTTTTTATCAGCGAG 8720
QY 1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
Db 8721 GACAGCATGCTCATCTCTTGACAGAGGACAGGTAAGGTAAATGGTTCAFCCTTCGATCG 8780
QY 1028 pPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
Db 8781 GTTCTGAGACAGGATATCTGTATTAACATAACAAGAACCCAGCTGATTTTACTACAATCAG 8840
QY 1048 rAsnAspAsnGlnTyrMetAspTyrArg 1057
Db 8841 TAAAGATAACAGTATATGGATTACCGC 8868
RESULT 4
US-10-425-114-4042
; Sequence 4042, Application US/10425114
; Publication No. US20040034988A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4042
; LENGTH: 871
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700347436_FLI
US-10-425-114-4042
Alignment Scores:
Pred. No.: 9,19e-37 Length: 871
Score: 491.50 Matches: 107
Percent Similarity: 65.79% Conservative: 18
Best Local Similarity: 56.32% Mismatches: 50
Query Match: 8.89% Indels: 15
DB: 13 Gaps: 6
US-09-721-114-2 (1-1057) x US-10-425-114-4042 (1-871)
QY 871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu-----SerGlnLeuHisAsnSerGln 888
Db 2 AACTACCAAGAGAGCTAATAGCGCGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAG 61
QY 889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
Db 62 TACACACAGCATGAGCATTTACATATGCTCACCGACACGTCATATGGAAGCCAA----- 115
QY 908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
Db 116 -----CCGCTACACTGGAAGACTGTCTCGGGTCAATTCACGACAGACTTGGC 166
QY 927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
Db 167 AGGCGCTTTACGCCCTCACCCCTCGTGTGGTGTGCTCGGTTCTATTCGTGACGAGGAGATC 226
QY 947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966
Db 227 GAAACTGTCTCTGGAACTGCGGCGCGAGTCTGGGTACAGACTAGTGTGATTTGAAGGG 286
QY 967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986
Db 287 ACAAGCGCGCTGATGCCCAACAGAGCAGGAGAACTACGAGACCTTCAGC----- 334


```

Db 1253 AAAAGGCAATAGCTAGCTCTATATATTTCTAAAGAGAAACGCAACCCAGTCT 1312
Qy 461 -----ThrGlySerValHisHisThrValAlaHisPro 471
Db 1313 GAGTCTTTCTAATTATGACTCAGAAATAGAAAAAGAGATAAAGACATGAGTAAATGGT 1372
Qy 472 AlaGlyAsnLeuSerAspLysValThrProThrAlaSer---ThrGlnHisAspAsp 490
Db 1373 GCTCCAGAACCCACAAAAGAAATCCAAATACAAAGATTTTGACTCTTCTGAGAT 1432
Qy 491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509
Db 1433 GAGAAACACAGCAAAAGAAAGAAATGATAATCAAGGGCACAAAATTTGAAGACCTCCAA 1492
Qy 510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db 1493 GAAGATCATCTGATGCTGTGAAGA---AAACAAGAGAGAGAGACTTCTCTTCAGCA 1549
Qy 530 LysGly-----Lys 532
Db 1550 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTTAAGAGATCGACTTCTTAAGAAG 1609
Qy 533 ThrHisSerAlaAlaSerThrLysThrGlyGlyGlnSerThrArgAsnGlyGlnAsnIle 552
Db 1610 CAGCAAGCAAGTGTCTCCACT-----GATGGTGTGCGATAAGCTTCTCGGAAAGAGCAG 1663
Qy 553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGlnAsnSerValLeuSerHis 572
Db 1664 AGTTTACTTCTTGGAGTTAGAAAGTCTCGAACTAAGAAAG-----AGCAAG 1717
Qy 573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
Db 1718 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTCGAGAGAA 1777
Qy 593 SerLeuProLysLysLys-----LysLysGlnLysLeuGlu 604
Db 1778 TTCCTAAGAAGACCGAGCGATGAACCTTCTGAAGATGATTAAGACAGCAAGAAAAG 1837
Qy 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db 1838 GGAAGTGAAGAAAAAAGAAA-----CCTTCAGACTTTAAGAAATGAAGAAAGTA 1882
Qy 625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp----- 641
Db 1883 ATTAAA-----ATGGAACAACAGTATGAATCTTCATCTGATGCGCATGAAAGTTACCT 1936
Qy 642 -----IleAsnArgIleGlnSerLysThrThr 650
Db 1937 GAGCGAGAGAAATTTGTTCATTTCTTAAGGCGATATAACAAATTAAGATGGAACACT 1996
Qy 651 AlaAspAspCysValIleValAlaLysAspGlySerAspTyrAlaSerSerVal 670
Db 1997 -----GATGGAGAAAAGAAAAGTAAAAAATA 2023
Qy 671 PheAspThrAsnSerGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690
Db 2024 AGAGATAAACTTCTAAAGAGAGATGAATATCTGATATGCTGAGAGTCAA--CAG 2081
Qy 691 GlyHisLeuAlaLeuThrThrGlnGlnSerProHisProGlnAsnPheGlnSerThrGln 710
Db 2082 GGA----- 2084
Qy 711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730
Db 2085 -----AAGAGATAGTTGTGACTCTT- 2105
Qy 731 PheSerHisAspAspGlnThrIleAlaGluAlaProThrGluHis--TrpGlyArgGly 750
Db 2106 -----CAGAGATAAAAGAGATGAAGATGGAGCATATGTTAGAGAGA 2147
Qy 750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765

```

```

Db 2148 AGAAAAGTGCAGAGTTGCTTGGAAAGAGTTCAAGGAGAGACAAGA 2193
RESULT 6
US-10-252-157-188
; Sequence 188, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Pearson, Mary
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 188
; LENGTH: 11167
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 261982.8
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 9131, 9138, 10638
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-188
Alignment Scores:
Pred. No.: 8,13e-10 Length: 11167
Score: 223.00 Matches: 172
Percent Similarity: 34.31% Conservative: 108
Best Local Similarity: 21.08% Mismatches: 312
Query Match: 4.04% Indels: 225
DB: 15 Gaps: 33
US-09-721-114-2 (1-1057) x US-10-252-157-188 (1-11167)
Qy 62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
Db 1425 AAGGCTCACTTCATTCGATTCGGAAGAGACTTAAATTCGAGTTTCGAGCGAGTGTCTGTA 1484
Qy 82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
Db 1485 ---ARCAAGAGAGAAAATACCAAGAGCATAAAGTCATAGATCTAGTCTTGAACACAAA 1541
Qy 101 -----AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
Db 1542 GCACGAAAGGAGAAAACCTTGTGCTTTGGAAAAGAGGATATTTCAAAGTCAGAGCT 1601
Qy 114 LysLeu----- 115
Db 1602 AAACCTTCAAGAAAACAGGTAGTAGTAGCAGCATGCATCAGAAATGTTCCAAAGAGGAA 1661
Qy 116 -----LysThrSerAspAsnGlyThrAlaPro 124
Db 1662 CAAAGAACAAATAAAAGTACCGGTGGTGAACATCTAGTAGAAAAGAGAACTT 1721
Qy 125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
Db 1722 CAATATGAACCTGCC-----AACACTCTGGAAGATTTAGACATGATATTGTGTCT 1772
Qy 145 SerThrPheValProAlaSerValGlySerGln-----LysValSerPro 159
Db 1773 -----GTTCTTCTCAGTCTCAGAGAGACATTTTGGAGAAATCTTGAGACTGTGTATG 1823
Qy 160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db 1824 GAAGTTTCAGAGTTCACTTGTATCATCAAGGGATGCGAGCAGTGAACCTGAACAAAGAGTG 1883
Qy 180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn-----GlyAlaAlaGlu 197

```

Db 1884 GAGGATTCATCTGTAATAATTAATTTCTTCAAAAGACACAGAGAGGTATTAATCA 1943
Qy 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db 1944 AAAATACAGCTAAGTAACAAAGAAATTA-----TATGTTAAACTC 1985
Qy 218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db 1986 ACTCCTGTTTCCCTTCTTAATTCCTCAATTAAGAGTGTGCTGATTTGTCAGGAAGTTCCACAA 2045
Qy 238 IleThrTrpHisIleGluValAsnGlyValaAspGlnProProSerThrProLysLeuSer 257
Db 2046 -----GATAAAGATGGCTATATAAAGTTGGTCTGAACCCCAAGTTAGAG 2090
Qy 258 GluValValLeuLysArgAsnGluAspGluAsnGly-----LysThrGluGluThr 274
Db 2091 AAATGTGACATTCGACAGGAAACAGTGAATATGAGCAATTTGGTTGAAATGAAGTTTCA 2150
Qy 275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db 2151 TTACTTTAGAGGAATCTGATCTCGAGATCCCGCTGAAGACTACACCTTGAGG 2210
Qy 295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
Db 2211 CGACCGACA-----GAAACTAACCTCTAATCAATTAATCA 2246
Qy 315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334
Db 2247 GATGAAGAA-----TGTATGAACAGTTTAGAG-----AAACAAAATCATCGTTCCA 2297
Qy 335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db 2298 GTGAGAAAAGGATAAGCGT----- 2318
Qy 355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal-----Arg 372
Db 2319 -----AATCTCTGACAGTGTATAGATAAATCCTAAGCCTAATAAATGGCCAAATCT 2372
Qy 373 LeuLeuSerGluIleLeuAsnAlaAsnGlnValGluAspSerArg-----SerAsp 389
Db 2373 AAGCAATCAGAGACTGGGATCAAAATTCAGATCTCTGATGAATGCTAGCAATCCTCAAA 2432
Qy 390 GluValHisArgGluAsnAlaAlaAspProCysGluAspArgSerThrIleProVal 409
Db 2433 GAGGTAGCAGGATGATGACATCTCTCTTCAGATACCTGATATTAATGAATTCATACA 2492
Qy 410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db 2493 AACCTAAGACTTGTATGATTTAAGACTCAG-----CCGGGGAAGAT----- 2537
Qy 430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp----- 446
Db 2538 -----GATAAAGGAAAAGGAAACGAAAAGTTCTACATCTGGCTCAGATTTTGATCT 2591
Qy 447 -----GlySerSerLeuMetAsnTyrLeuAsnGlyLysLysLysArg----- 460
Db 2592 AAAAAGGGCAATCAGCTAGAGCTCTATATTTCTTAAAGAAAGACGAAACCCAGTCT 2651
Qy 461 -----ThrGlySerValHisHisThrValAlaHisPro 471
Db 2652 GAGTCTTCTAATTAGTACTCAGAAATAGAAAAGAGATAAAGAGCATGAGTAAATTTGGT 2711
Qy 472 AlaGlyAsnLeuSerAsnLysValThrProThrAlaSer-----ThrGlnHisAspAsp 490
Db 2712 GCTGCCAGAACCCAAAAAGAAATTCCAAATACAAAAGATTTTGACTCTTCTTGAGAT 2771
Qy 491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr-----AspValCysGln 509
Db 2772 GAGAAACACAGCAAAAAGGAATGATATCAAGGGCACAATAAATTTGAAGACTCCACAA 2831
Qy 510 HisValSerGluLeuSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db 510 HisValSerGluLeuSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529

Db 2832 GAAGGATCATCTGATGATGCTGAAAGA-----AAACAAGAGAGAGAGACTTTCTTCTTCAGCA 2888
Qy 530 LysGly-----Lys 532
Db 2889 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGAATCTTCTTAAGAAG 2948
Qy 533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552
Db 2949 CAGCAAGCAGAGTCTTCCACT-----GATGGTGTGATAGACTTTCTGGGAAGAGCAG 3002
Qy 553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572
Db 3003 AGTTTACTTCTTGGAGTTAGAAAAGTTGCTGAAACTTAAAGAAAAG-----ACCAAG 3056
Qy 573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
Db 3057 CATCTCAAAACCAAACATGTAAAAAGTACAGGATGGCTTATCTGATATTGACAGAAA 3116
Qy 593 SerLeuProLysLysLys-----LysLysGlnLysLeuGlu 604
Db 3117 TTCCTAAAGAAAGACACAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGACGCAAAAAG 3176
Qy 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db 3177 GGAATCTGAAGAAAAGAAAGAA-----CCTTCAGACTTTAAGAAAAGTA 3221
Qy 625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp----- 641
Db 3222 ATTAAA-----ATGGAACAACAGTATGAATCTTCTATCTGATGGCACTGAAAAGTTACCT 3275
Qy 642 -----IleAsnArgIleGlnSerLysThrThr 650
Db 3276 GAGCGAGAGAATAATTGTCAATTTCTTAAGGCATATAAAACAAATTAAGAAATGGAACAAC 3335
Qy 651 AlaAspAspCysValIleValAlaLysAspGlySerAspTyrAlaSerSerVal 670
Db 3336 -----GATGGGAAAAGAAAGTAAAAAATA 3362
Qy 671 PheAspThrAsnSerGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690
Db 3363 AGAGATAAAACTTCTTAAAGAAAGGATGAATATCTGATTTATGCTGCAAGAGTCAA--CAG 3420
Qy 691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710
Db 3421 GGA----- 3423
Qy 711 GluGlnGlnThrHisLeuArgMetGluLeuMetValThrIleAlaProThrGluHis-TripGlyArg 730
Db 3424 -----AAGGAGATAGTTGTGACTCT----- 3444
Qy 731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TripGlyArg 750
Db 3445 -----CAGAGATAAAAGAGTAAGTAAGTGGAGCATATGTTAGAGAGA 3486
Qy 750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db 3487 AGAAAAGGTGCAAGTTGCTTGGAAAAGAGTTCAAGCAAGAGACAAGA 3532
RESULT 7
US-09-294-093B-2759
; Sequence 2759, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalquidi, Raghunath, V.
; APPLICANT: Ito, Laura Y
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207

```

; SOFTWARE: PERL Program
; SEQ ID NO 2759
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No. US20010051335A1 700346612H1
; NAME/KEY: unsure
; LOCATION: 28, 123, 150, 179, 191, 264
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-2759

Alignment Scores:
Pred. No.: 4,87e-12 Length: 279
Score: 222.50 Matches: 55
Percent Similarity: 64.50% Conservative: 7
Best Local Similarity: 57.25% Mismatches: 25
Query Match: 4.03% Indels: 10
DB: 9 Gaps: 4

US-09-721-114-2 (1-1057) x US-09-294-093B-2759 (1-279)
Qy 871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu-----SerGlnLeuHisAsnSerGln 888
Db 2 AACTACCAAGAGCTAATAAGCGGCANATGGAGCTTCAACGCAAGCTCTCGGCTCGCAG 61

Qy 889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSerThr 907
Db 62 TACACACAGCATGACCAATACCAATCGCTCACGAGCAGCTCATATGGAAGCCAA----- 115

Qy 908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
Db 116 -----CCGCTGAACTGGAAGACTTGCTCGGCGCTCATTCACCAAGACTTGGCG 166

Qy 927 ArgProLeuArgProHisProArgValGlyValLeuGlyValSerLeuLeuGlyLysGluIle 946
Db 167 AGGCGTTTACGCNCTCACCTCGTGCNTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 226

Qy 947 AlaMetTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGly 962
Db 227 GCAAACTGGTCTGG-GAATGCGGCGCGCAGTCTGGGTANAGACTAGGT 273

RESULT 8
US-09-864-761-19041/c
; Sequence 19041, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeocelca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19041
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000511.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AW867076.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUATE 5.00e-63
; OTHER INFORMATION: NT HIT: A3001609.1, EVALUATE 6.50e-01
US-09-864-761-19041

Alignment Scores:
Pred. No.: 3,59e-10 Length: 3953
Score: 220.00 Matches: 153
Percent Similarity: 35.32% Conservative: 125
Best Local Similarity: 19.44% Mismatches: 311
Query Match: 3.98% Indels: 196
DB: 9 Gaps: 38

US-09-721-114-2 (1-1057) x US-09-864-761-19041 (1-3953)
Qy 113 AspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsn 132
Db 2346 GACAGGACTCTTGGCCAAATGAGAGACCCGCCATCTCTAGCAGAGCCTACAGAAAT 2287

Qy 133 GlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerVal 152
Db 2286 GGACAA-----AGGACCCCAATTGCCAATGAGAAGACC 2254

Qy 153 GlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer 172
Db 2253 ACATCATCTTCAGCAGAGCCTACAGAACACGAGAAGAGGACTCCACTGGCCCAATGAGAAC 2194

Qy 173 ThrLeuProLysSerValGlnGlyAsnAspSerLysCysAsnAlaProSerGlyLys 192
Db 2193 ACCACACCA---TCCCGCGCAGAGCCTACAGAAATAGA-----GAA 2155

Qy 193 AsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGln 212
Db 2154 AGGACAGCCCAATGAGAACACACCATCCCA-----GCAGGGCCTACAGAA 2107

Qy 213 AsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeu 232

```

Db 2106 AATGAGAAATGACAGCCAAC-----GAGAGACCACTATTCCGAGCAGAGCCT 2056
Qy 233 ProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSer 252
Db 2055 ACAGAAATAGAGAAAGCAGACCAATGAGAGAGACCACTATCCCGAGCAGCCTACA 1996
Qy 253 -----ThrProLysLeuSerGluValValLeuLysArgAsnGluAspGlu 267
Db 1995 GAAATGCAAAAGAGACCCATTTGCCAATGAGAAACCACTATCCCGAGCAGAGCCT 1936
Qy 268 AsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro--- 286
Db 1935 ACAGAACGAGAGAAAGCCCACTGGCCCAATGAG---AACACCACTATCCCGAGCA 1879
Qy 287 AsnProMetSerGlyLysGlu----- 293
Db 1878 GAGCCTACAGAAATAGAGAAAGCAGACCCCAATGAGAGACCCACACCATTTCCGAGCAGAG 1819
Qy 294 -----ArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspPro--- 307
Db 1818 CCTACAGAAATAGAGAAAGCAGACCAATGAG---AACACCACTATCCCGAGCAGAG 1762
Qy 308 -----LysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
Db 1761 CCTACAGAAATGAGACAGCAGCCTCCATTTGCCAATGAGAGACCCACACCATCTCTAGCA 1702
Qy 321 GluProCysGluGluValValLeuLysArgSerLysSerLysArgLysThrAspLys 340
Db 1701 GAGCCTACAGAAATGAG---AAAGGACCCATTTGCCAATGAGAGACCACTACA 1648
Qy 341 LysLeuMetLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
Db 1647 TCCTCAGCAGAGCCTACAGAACGCGAGAAAGG----- 1615
Qy 361 AlalysLeuCysArgLysProLysValArgLeuLeuSerGluIleIleAsnAla 380
Db 1614 -----ACTCCACTGGCCCAATGAG 1597
Qy 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAla----- 396
Db 1596 AACACCATCATCCCGAGCAGAGCCTACAGAAATAGAGAGAGCAGCAGCAATGAGAG 1537
Qy 397 -----AlaAspProCysGluAspArgSer-----Thr 406
Db 1536 ACCACCAATTTCCCGAGCAGAGCCTACAGAAATAGAGAAAGCAGCAGCAATGAGAGACC 1477
Qy 407 IleProValProMetGluValSerMetAspIleProValSerAsnHis-----ThrVal 424
Db 1476 ACACCATTCGAGCAGAG-----CCTACAGAAATAGAGAGATGGACAGCC 1432
Qy 425 GlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysLysSerAspValVal 444
Db 1431 AATGAGAACACCACTATCCCGAGCAGAGCCTACAGAAAT----- 1390
Qy 445 AspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerVal 464
Db 1389 GAAGAAATGACCCCATTTGGCCCAAT-----GAGAGACCACTATCC--- 1348
Qy 465 HisHisThrValAlaHisProAlaGlyAsn-----LeuSerAsnLysLys 479
Db 1347 -----CCAGCAGAGCCTACAGAAATGAGAGAGAGCCCAATTTACCAATGAGAG 1297
Qy 480 ValThrProThrAlaSer-----ThrGlnHisAspAspGluAsnAspThrGluAsnGly 497
Db 1296 ACCACCATCTCCGAGCAGAGCCTACAGAAATGAGAGAGAGCCCGCCCAATGAG 1237
Qy 498 LeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGln 517
Db 1236 ATC---ACCACCATCCCGAGCAGAGCCTACAGAACATGAGAGAGATAGGCCAATGAG 1180
Qy 518 ArgCys-----SerSerLysGlyLysThrAlaGlyLeuSerLys 530
Db 1179 AAGGCCACACCATCCCGAGCAAGCCCTACAGAACATGAGAGAAAGACA---GTCAATGAG 1123

RESULT 9

US-10-029-386-20602/c
; Sequence 20602, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

Qy 531 GlyLysThrHisSerAlaAlaSerThrLysTyrGlyGly----- 543
Db 1122 GACACCACTCATCTCCAGCAGAGCCTACAGAAATGAGAGAGACCCCACTGGCCCAAT 1063
Qy 544 -----GluSerThrArgAsnGlyGlnAsnIleHisValLeu 555
Db 1062 GAGAACACCACTATCCCGAGAGAGTCTACAGAACTGGAGAGAGAGCAGCCCAATGAG 1003
Qy 556 SerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys 575
Db 1002 AAGACCACTCATCTCCCGAGCAGAGCCTACAGAACTGGAGAGAGCAGCCCAAT 943
Qy 576 -----ValSerProAlaGluHisAspIleGlnIleMetSerAsp 588
Db 942 GAGAGACCATACCATCTCCCGAGAGAGCCTACAGAACTGGAGAGAGCAGCCCAAT 886
Qy 589 LeuHisGluGlnSerLeuProLysLysLysLys-----GlnLysLeuGlu 604
Db 885 GCCAATGAGAGACCACTATCCCGAGAGAGCCTACAGAACTGGAGAGAGAGCAGTACA 826
Qy 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db 825 TTGGCCCAATGAGAGAGATCACTATCCCGAGAGAGCCT-----ACAGAACATGGA 775
Qy 625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArg 644
Db 774 GCAAAATACCTGCGCCCAATGAGAGATCACTATCCCGAGAGAGCCTACAGAACT 715
Qy 645 IleGlnSerLysThrThrAlaAspAspCysValIleValAlaLysAspGlySer 664
Db 714 ---GGAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 658
Qy 665 AspTyrAlaSerSerValPheAspThrAsn----- 674
Db 657 GAACATAGAGTAGGCTACATCAGCCATGTGATCAGACAGAGCAGCCAGCAGCCTATA 598
Qy 675 -----SerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnHis 692
Db 597 AACATGCAAAAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 553
Qy 693 LeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGln 712
Db 552 -----GTACAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 499
Qy 713 GlnThrHisLeuArgMetGluLeuMetValThrIleAlaAlaSerSerProLeuPheSer 732
Db 498 ACCACA-----AGAACCCCGAGAGAGCCTACGCTATCTACAGAGAGAGCAGCAGC 445
Qy 733 HisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAla 752
Db 444 AAAGGAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 394
Qy 753 LysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsn-----SerProAlaAla 770
Db 393 ACCACACTGACCACTGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 334
Qy 771 ThrCysGlyAla-----GlnPheArgPro-----GlyIleGlnAlaVal 783
Db 333 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 274
Qy 784 AspLeuThrSerThrHisVal 790
Db 273 ACTACTACT 253

; TITLE OF INVENTION: HUMAN GEMCMB-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ACOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annotax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20602
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB023048.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 6.00e-63
; OTHER INFORMATION: EST HUMAN HIT: AW867076.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: AY026090.1, EVALUE 2.50e-01
US-10-029-386-20602

Alignment Scores:
Pred. No.: 3,59e-10 Length: 3953
Score: 220.00 Matches: 153
Percent Similarity: 35.32% Conservative: 125
Best Local Similarity: 19.44% Mismatches: 311
Query Match: 3.98% Indels: 198
DB: 15 Gaps: 38

US-09-721-114-2 (1-1057) x US-10-029-386-20602 (1-3953)
Qy 113 AspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsn 132
Db 2346 GACAGGACTCTTTGGCCATGAGAACACCGCCGCTCTAGCAGAGCTCAGAAAT 2287
Qy 133 GlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerVal 152
Db 2286 GGCACAA-----AGGACCCCATTTGGCCATGAGAACGACC 2254
Qy 153 GlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer 172
Db 2253 ACATCATCTCCAGCAGAGCTCAGAACACGAGAAAGGAGCTCCATGGCCATGAGAAC 2194
Qy 173 ThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLys 192
Db 2193 ACCACACCA---TCCCGCGCAGAGCTCAGAACAAATAGA-----GAA 2155
Qy 193 AsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGln 212
Db 2154 AGGACGCCAATGAGAACACACCACTCCCA-----GACGGGCTCAGAA 2107
Qy 213 AsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeu 232
Db 2106 AATAGAGAAATGACAGCAAC-----GAGAGACCACTATTTCCAGCAGAGCCT 2056
Qy 233 ProGluValProGlnIleThrThrPheIleGluValAsnGlyAlaAspGlnProSer 252
Db 2055 ACAGAAATAGAGAAAGGACGCAATGAGAACCAACCATCATCTCCAGCAGAGCCTACA 1996
Qy 253 -----ThrProLysLeuSerGluValValLysLysArgAsnGluAspGlu 267
Db 1995 GAAATCGACAAAGGACCCCATTTGGCCATGAGAAACCACTATCTCCAGCAGAGCCT 1936
Qy 268 AsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro---- 286
Db 1935 ACAGACACGAGAAAGGACCCCACTGGCCCAATGAG---AACACCACTATCTCCAGCA 1879
Qy 287 AsnProMetSerGlyLysGlu----- 293
Db 1878 GAGCCTACAGAAATAGAGAAAGGACGAGCCCAATGAGAACCACTATCTCCAGCAGAG 1819

294 -----ArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspPro----- 307
Db 1818 CCTACAGAAATAGAGAAAGGACGCAATGAG---AACACCACTATCTCCAGCAGAG 1762
Qy 308 -----LysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
Db 1761 CCTACAGAAATGAGACGAGCTCCATTTGGCCATGAGAACCACTATCTCTAGCA 1702
Qy 321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
Db 1701 GAGCCTACAGAAATGGA-----AAAGGACCCCATTTGGCCATGAGAACCACTATCA 1648
Qy 341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
Db 1647 TCCTCAGCAGAGCTCAGAACACGCGAGAAAG----- 1615
Qy 361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
Db 1614 -----ACTCACTGGCCCAATGAG 1597
Qy 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAla----- 396
Db 1596 AACACCACTATCTCCAGCAGAGCTCAGAAATAGAGAAAGGACGAGCAATGAGAG 1537
Qy 397 -----AlaAspProCysGluAspAspArgSer-----Thr 406
Db 1536 ACCACAAATTCACGAGAGCTCAGAAATAGAGAAAGCAGACCAATGAGAGACCC 1477
Qy 407 IleProValProMetGluValSerMetAspIleProValSerAsnHis-----ThrVal 424
Db 1476 ACACCAATTCACGAGAG-----CCTACAGAAATAGAGAAATGAGAGACCC 1432
Qy 425 GlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValVal 444
Db 1431 AATGAGAACACCACTATCTCCAGCAGAGCTCAGAACAT----- 1390
Qy 445 AspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArgThrGlySerVal 464
Db 1389 GAAGAATGACCCCATTTGGCCAT-----GAGAGACCACTATCTCC 1348
Qy 465 HisHisThrValAlaHisProAlaGlyAsn-----LeuSerAsnLysLys 479
Db 1347 -----CCAGCAGAGCTCAGAAATGAGAAAGGACCCCATTTACCAATGAGAG 1297
Qy 480 ValThrProThrAlaSer-----ThrGlnHisAspAspGluAsnAspThrGluAsnGly 497
Db 1296 ACCACCACTATCTCAGCAGAGCTCAGAACATGAGAAAGGACCCCACTGGCCCAATGAG 1237
Qy 498 LeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGln 517
Db 1236 ATC---ACCACCACTCCGAGCAGAGCTCAGAACATGAGAAAGGAGGATGAGCAATGAG 1180
Qy 518 ArgCys-----SerSerLysGlyLysThrAlaGlyLeuSerLys 530
Db 1179 AAGGCCACACCACTCCCGAGCAAGCTCAGAACATGAGAAACGACA---GTCAATGAG 1123
Qy 531 GlyLysThrHisSerAlaHisSerThrLysTyrGlyGly----- 543
Db 1122 GACACCACTATCTCAGCAGAGCTCAGAAATGAGAAAGGACCCCACTGGCCCAAT 1063
Qy 544 -----GluSerThrArgAsnGlyGlnAsnIleHisValLeu 555
Db 1062 GAGAACACCACTCCCGAGCAAGCTCAGAACATGAGAAAGGACGAGCAATGAG 1003
Qy 556 SerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys 575
Db 1002 AAGACCACTATCTCCAGCAGAGCTCAGAACATGAGAAAGGACCACTATCTCCAGCA 943
Qy 576 -----ValSerProAlaGluHisAspIleGlnIleMetSerAsp 588
Db 942 GAGAGACCACTATCTCCAGCAAGCTCAGAACACGAA---GAATGAGCCCATCG 886
Qy 589 LeuHisGluGlnSerLeuProLysLysLysLysLys-----GlnLysLeuGlu 604


```

Db 885 GCATGAGAACACACACATCCCGTAGTAAGCCTACAGAACATGAGAGAAAGACTACA 826
Qy 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeu 624
Db 825 TTGGCCAAATGAGAAGATCACTATCCCGAAGGGCCT-----ACAGAACATGGA 775
Qy 625 AlaIysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArg 644
Db 774 GCAAAACTAGCTGGCCAAATGAGAAGATCACTATCCCGAAGGGCCT-----ACAGAACATGGA 715
Qy 645 IleGlnSerLysThrThrAlaAspAspCysValIleValAlaAlaLysAspGlySer 664
Db 714 --GGAGAAAGGACCATACCCCAATGACAAGATCACCTCATCTGCGAGCAGGTCTACA 658
Qy 665 AspTyrAlaSerSerValPheAspThrAsn----- 674
Db 657 GAACATAGATAGGGCTACATCAGCCCAATGTGATCACACGCCGCCCGAGAGCCCTATA 598
Qy 675 -----SerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHis 692
Db 597 AAACATGCAAAAGGACCATCTGGCCCATGAGAAGATGACAAA----- 553
Qy 693 LeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGln 712
Db 552 -----GTCAGAGAAAGTCCACAGAACACCCAGAAAGACCCAGTCAACACAGAGAAA 499
Qy 713 GlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSer 732
Db 498 ACCACA-----AGAACCCAGAAAGCCTAAGCTATCTACAGAGAACCATATGCACC 445
Qy 733 HisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAla 752
Db 444 AAAGGAGAAACACACACAGTCCACAGAAAGCCTACAGAAAGCCTGGGGAAAC----- 394
Qy 753 LysLysLeuThrTrpGluGlnPheLeuAlaThrThrArgAsn-----SerProAlaAla 770
Db 393 ACCACATGACCTAGACCATTAAGCCCGGAGTAAAGTCCAGAGAACCCAGAGAAAC 334
Qy 771 ThrCysGlyAla-----GlnPheArgPro-----GlyIleGlnAlaVal 783
Db 333 ACAGCAGCAGTCACAAAGACTATAAAACCTTCAGTCAAGGTACAGAGGAGCAAAATCTCTC 274
Qy 784 AspLeuThrSerThrHisVal 790
Db 273 ACTACTACTCTCTCTCATCTA 253

RESULT 10
US-10-295-027-427
; Sequence 427, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-295-027-427

Alignment Scores:
Pred. No.: 4,15e-10 Length: 4383
Score: 220.00 Matches: 153
Percent Similarity: 35.32% Conservative: 125
Best Local Similarity: 19.44% Mismatches: 311
Query Match: 3.98% Indels: 198
DB: 16 Gaps: 38

US-09-721-114-2 (1-1057) x US-10-295-027-427 (1-4383)
Qy 113 AspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsn 132
Db 1891 GACAGAGCTCTTTGGCCCAATGAGAAGACCGCCATCTCTAGCAGAGCTACAGAAAT 1950
Qy 133 GlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerVal 152
Db 1951 GGACAA-----AGACCCCATTTGCCCAATGAGAAGACC 1983
Qy 153 GlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer 172
Db 1984 ACATCATCTCTCAGCAGAGCTTACAGACACAGAGAAAGGAGCTCCACTGGCCCAATGAGA 2043
Qy 173 ThrLeuProLysSerValGlnGlyAsnAspSerLysCysAsnAlaProSerGlyLys 192
Db 2044 ACCACACCA---TCCCGGCGAGAGCTTACAGAAATAGA-----GAA 2082
Qy 193 AsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGln 212
Db 2083 AGGACAGCCCAATGAGACACACCATCTCCCA-----GCAGGGGCTACAGAA 2130
Qy 213 AsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeu 232
Db 2131 AATAGAGAAATGACAGCCAAAC-----GAGAAAGACCACTATTTCCCGCAGAGCT 2181
Qy 233 ProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProSer 252
Db 2182 ACAGAAATAGAGAGAGGACAGCCATGAGAGACCACTATCTCCCGCAGAGCTTACA 2241
Qy 253 -----ThrProLysLeuSerGluValValLeuLysArgAsnGluAspGlu 267
Db 2242 GAAATGAGACAAAGGACCCCATTTGCCAATGAGAAACCCACATCATCTCCCGCAGAGCT 2301
Qy 268 AsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 2302 ACAGAACCGGAGAAAGGACCCCATCTGGCCCAATGAG---AACACCACTATCTCCCGC 2358
Qy 287 AsnProMetSerGlyLysGlu----- 293
Db 2359 GAGCCTTACAGAAATAGAGAGAGGACAGCCCAATGAGAGACCAACCATCTCCCGCAGAG 2418

```

```
QY 294 -----ArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspPro----- 307
Db 2419 CCTACAGAAAATAGAGAGACGACCAATGAG---AACACCAACCATCCCGACGACAG 2475
QY 308 -----LysProValSerGlyGlnLysCysGluGlnLeuCysAsn 320
Db 2476 CCTACAGAAAATGAGACGACGACTCCATGGCCATGAGAGACCAACCATCTCTAGCA 2535
QY 321 GluProCysGluGluValLeuLysArgSerLysSerLysArgLysThrAspLys 340
Db 2536 GAGCTACAGAAAATGGA-----AAAAGACCCCATTTGCCAATGAGAGACCAATCA 2589
QY 341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
Db 2590 TCCTCAGCAGGCTACGAGAACCGCGAAGG----- 2622
QY 361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleLeuAsnAla 380
Db 2623 -----ACTCCACTGGCCCAATGAG 2640
QY 381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAla----- 396
Db 2641 AACACCAATCATCCCGACGAGCGCTACAGAAAATAGAGAAAGGACAGCCCAATGAGAAG 2700
QY 397 -----AlaAspProCysGluAspArgSer-----Thr 406
Db 2701 ACCACCAATTCCTCAGCAGCGCTACAGAAAATAGAGAAAGGACGCAATGAGAGACC 2760
QY 407 IleProValProMetGluValSerMetAspIleProValSerLysHis-----ThrVal 424
Db 2761 ACACCATTCCTCAGCAGAG-----CCTACAGAAAATAGAGAAATGAGAGAGCC 2805
QY 425 GlyGluAspGlyLeuLysSerLysAsnLysThrLysArgLysTyrSerAspValVal 444
Db 2806 AATGAGAACACCAACACTATCCCGACGAGCGCTACAGAACT----- 2847
QY 445 AspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerVal 464
Db 2848 GAAGAAATGACCCCATTTGGCCAAAT-----GAGAGAGACCAACTATCC--- 2889
QY 465 HisHisThrValAlaHisProAlaGlyAsn-----LeuSerAsnLysLys 479
Db 2890 -----CCAGCAGAGCGCTACAGAAAATGAGAAAGGACCCCATTTACCAATGAGAAG 2940
QY 480 ValThrProThrAlaSer-----ThrGlnHisAspAspGluAsnAspThrGluAsnGly 497
Db 2941 ACCACCACTCTCAGCAGAGCGCTACAGAACTGGAGAAAGGCCCACTGGCCCAATGAG 3000
QY 498 LeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGln 517
Db 3001 ATC---ACCACACCATCCCGACGAGCGCTACAGAACTGGAGAAAGGATAGCCCAATGAG 3057
QY 518 ArgCys-----SerSerLysGlyLysThrAlaGlyLeuSerLys 530
Db 3058 AAGGCCACACCATCCCGACGAAAGCGCTACAGAACTGGAGAAAGGACCA---GTCNAATGAG 3114
QY 531 GlyLysThrHisSerAlaAlaSerThrLysTyrGlyGly----- 543
Db 3115 GACACCAACCATCTCAGCAGAGCGCTACAGAAAATGAGAAAGGACCCCACTGGCCCAAT 3174
QY 544 -----GluSerThrArgAsnGlyGlnAsnIleHisValLeu 555
Db 3175 GAGAACCAACCAACATCCCGACGAGCTCTACAGAACTGGAGAAAGGACAGCCCAATGAG 3234
QY 556 SerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys 575
Db 3235 AAGACCAACCATCCCGACGAGCGCTACAGAACTGGAGAAAGGACCACTAGCCCAAT 3294
QY 576 -----ValSerProAlaGluHisAspIleGlnIleMetSerAsp 588
Db 3295 GAGAAAGACCATACCATCTCCAGCAAGCGCTACAGAACCGAA---GAATGAGCCCATCG 3351
QY 589 LeuHisGluGlnSerLeuProLysLysLysLysLys-----GlnLysLeuGlu 604
```

```
Db 3352 GCCAATGAGAACACCAACCATCCCGAGTAAAGCCCTACAGAACATGAGAGAAAGACTTACA 3411
QY 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db 3412 TTGGCCCAATGAGAGATGATCACACTATCCCGAGAGGCGCT-----ACAGAACATGGA 3462
QY 625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArg 644
Db 3463 GCAAAAACACTCGTCGGCCCAATGAGAGATGATCACACCATCCCTAGCAAAAGCCTCAGAACAT 3522
QY 645 IleGlnSerLysThrAlaAspAspCysValIleValAlaAlaLysAspGlySer 664
Db 3523 ---GGAGAAAGGACCACTACCCCAATGACAGATCACCTCATCTCGACGAGATCTTACA 3579
QY 665 AspTyrAlaSerSerValPheAspThrAsn----- 674
Db 3580 GAACATGAGATAGGCGCTACATCAGCCCAATGTGATCACACAGCCCGCCAGAGCCCTATA 3639
QY 675 -----SerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHis 692
Db 3640 AACATGCAAAAGGACCACTTGGCCCATGAGAGATGACACAA----- 3684
QY 693 LeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGln 712
Db 3685 ---GTCACAGAAAAGTCCACAGAACACCCAGAAAAGACCAACGTCACACAGAGAAA 3738
QY 713 GlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSer 732
Db 3739 ACCACA-----AGAACCCCGAGAAAGCGCTATCTATCTCAGAGAGACCATATGCACC 3792
QY 733 HisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAla 752
Db 3793 AAGGGGAAAAACACACCACTCCCGAGAAAGCGCTACAGAAAACCTGGGGAAAC----- 3843
QY 753 LysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsn-----SerProAlaAla 770
Db 3844 ACCACACTCACCCTGAGACCACTAAAGGCCCGCCAGTAAAGTCCACAGAAACCCAGAAAA 3903
QY 771 ThrCysGlyAla-----GlnPheArgPro-----GlyIleGlnAlaVal 783
Db 3904 ACAGCAGCAGTCCAGAAAGACTATAAACCTTCAGTCAAGCTCAGAGAGACAAATCTCTC 3963
QY 784 AspLeuThrSerThrHisVal 790
Db 3964 ACTACTACCTCTCTCTCATCTA 3984

RESULT 11
US-10-120-988-286
; Sequence 286, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Felyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt.PL_genes Version 2.0
; SEQ ID NO 286
; LENGTH: 6745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

QY	297	ValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGlu	316
DB	3850	ACATTACATAAAAGTTAAG---ACGACTAGTAGCTTTCTGTTCTGTTGATGAGTAT	3906
QY	317	GlnIleCysAsnGluPro-CysGluGluValValLysArgSerSerLysSerLys	335
DB	3907	GTAAAATGTCGGAGTGTCTCAATATATTACACTCTACCGAGGAACCCAGCAAAA	3966
QY	335	-----	335
DB	3967	FTCTGTAACTCCTTCAACAGTATACACAAATATAATTAATTACTATAGAACTCA	4026
QY	336	-----ArgLysThrAspLysLeuMet	343
DB	4027	GTGGAGCTGAACAATTTCTTAACGCTTTAGAAAGACAAACAGAAATATTTCAC	4086
QY	344	-LysLysGlnGlnHisSerLysArgThrAlaGlnAlaAspValSerAspAlaLys	363
DB	4087	GAGCAGCTCAGAACACCTTCATGTGAAATCTAAAGATGT-----CCGTCA	4137
QY	363	uCysArgLysProLysValArgLeuLeuSerGluIleAlaAsnAlaAsnGlnVa	383
DB	4138	GATCAGCCTCACCACCTGAAATATGCTGCTCCGATTATCA-AAATAGGGGCCC	4196
QY	383	lGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGlu	403
DB	4197	AGCTCCTACATTACAGGAATGCTTCTGTTGAGGAGCT-----	4236
QY	403	parGSerThrIleProValProMetGluValSerMetAspIleProValSerAsn	423
DB	4237	-----GTTCTCTTCCTGAGAGGAATCTAAAGCTAGAGATTTTTCAGATA	4286
QY	423	rValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTySer	443
DB	4287	TTTAGTAAACACCTCTAGGTGATTTCAGAAACCAAGAGGAAGA-----	4332
QY	443	lValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArgThrGly	463
DB	4333	-----GCCAAAGATTCCAAAGTGAAC	4355
QY	463	rValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThr	483
DB	4356	CCTGCATACCTTCATTGCTTCAGAGAAAAAATGTATCCGAGAAAAA---T	4412
QY	483	rAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr	503
DB	4413	TTGTCAACCAATCCATTAAATTCAGTAACAGTGTCTCTCTTCCAGCT---C	4469
QY	503	sLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSer	523
DB	4470	AGAGTATATATGGAAATTCACAACTAGAGAAGTCTCTGGAGTGTACAGGAGT	4529
QY	523	YlysThr-----AlaGlyLeuSerLysGlyLysThrHisSerAlaLase	538
DB	4530	TAGAGCCATTCCATTCTAGGAAGTGTCCCGAAGATGCACATCCACAGC	4589
QY	538	rThrLysTyGly-----GlyGluSerThrArgAsnGlyGlnAsnIleHisVal	556
DB	4590	TGTAGTGTATGCTCCAGTGGATCAGCCTAGGAGGACAGAGG-----GA	4637
QY	556	rAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAla	575
DB	4638	CATTGGAAACCACTGCCAAAAATGACT---AATAAAACACTTCTCTCACTCA	4694
QY	576	-----ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu	591
DB	4695	AGTCTTTGGCTTCTCCAGCATTGCTAAACTACAGCTTGTGTGAGGACTCAGT	4754
QY	592	-----GlnSerLeuProLysLys-----LysLysLysGlnTy	602
DB	4755	TGAACCAAACTTAGAGAGTCTGCAGTCTGAACCAAGAGAAATTAACCTCA	4814
QY	602	sLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle	622

LOCATION: (1)...(6147)

US-10-120-988-286

Alignment Scores:

Pred. No.: 9,51e-10 Length: 6745
 Score: 219.00 Matches: 213
 Percent Similarity: 33.71% Conservative: 172
 Best Local Similarity: 18.65% Mismatches: 410
 Query Match: 3.96% Indels: 348
 Gaps: 16

US-09-721-114-2 (1-1057) x US-10-120-988-286 (1-6745)

QY	83	HisAspGln-----LysLysCysAspGluHisLysLysSerSerProPheSer	99
DB	3034	CATGATGAAGGTGTGATGTCATTAAGCCATTTCACACTCTCTTTTAGGAATGA	3093
QY	100	ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThr	119
DB	3094	AAAGGAAAAATAGGCATCAT-----ATATCTGTATTGAAAGTTTAAAGCA	3147
QY	120	AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGly	139
DB	3148	AGTATATCAGTACCC-----ACCAGTGAATCAGAGGC	3180
QY	140	IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal	158
DB	3181	CTC-----ATTGAAGCAATCAAGCAATTCCTCAATGTTCTGAA	3219
QY	159	ProSerThrGlnSerSer-----GlnGlyLys	167
DB	3220	CTTGACACAATTTATTGTACCTTCCCAAGAAATCAAGCAGTTTCTCATACAT	3279
QY	168	AsnAlaAspArgSerThrLeuProLysSerValGlnGluGly	181
DB	3280	CAGTCAGGAAGTAATAATGCTGCTTCATTGAGGAATGGGCCACCTCCCTTCCA	3339
QY	182	-----AsnAspSerLysCysAsnAla	188
DB	3340	AAAAATAATGGAAGATGCAATGGGAATCTATATGTTAAACAAATTTAGTCC	3399
QY	189	ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAsp	208
DB	3400	CCTGAGTCAGGAATGAATGTTCCAAAGTCTTTCAGACTCAGCC-----CT	3450
QY	209	GlyProAlaGlnAsnTyAspValAlaAlaAsnValSerGluAspAsnThrSer	228
DB	3451	GCACCTGAAGCCACAGAGAGAATGACAAATGTAAAAAGCAGTGTCTTCTCG	3510
QY	229	ValGlyAlaLeuPro-----GluValProGlnIleThrTrpHis	241
DB	3511	AAAGGACCACTTCCATTCTCATCAAGGCTATGTCATGCTCCCTCAGGGAGCC	3570
QY	242	IleGlu-----ValAsnGlyAlaAspGlnProPro	251
DB	3571	GCCTCACTGGNAGAGGAGGAGAAAAAGCCATTGACCTCAGGCATGGATGCTT	3630
QY	252	SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGly	271
DB	3631	CTAACACCAAGGGCTTGGGAGAGAATCATTAGCCCTGTGGAAGTGAATCAT	3684
QY	272	GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys	284
DB	3685	-----GTTAGAGATTGTTCTTTTACCAAAAGACAAACCAAGGAAAC	3729
QY	284	-----	284
DB	3730	TTCCAAGAAATACACTGAGAAAGGGTAAAAATGGCTGCTCCAGGAGAGTGTAT	3789
QY	285	-----AspProAsnPro-----MetSerGlyLysGluArgAspGln	296
DB	3790	CTTTCAAATGAAGACCTTTTACCTTTGCTCAGACTTGTCCAGGAAGAAGCTG	3849

Db 4815 GGCAATATGACAGAGCAGGAG 4839
Qy 622 uLeuLeuAlaLysGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp11 642
Db 4840 -----GCTGAAGATGAATGAGAGAGTCACTGGGATCACTTCTCCTGAAG 4892
Qy 642 eAsnArgGlnSerLysThrAlaAspAspCysValIleValAlaLysAs 662
Db 4893 AAACAAA-----AATAAAACCACTTGGATGAC-----CTAGTAAAGGGGAAAA 4937
Qy 662 pGlySerAspTyr-----AlaSerSerValPheAspTh 673
Db 4938 TAGATCTTCGTAAACACAGATTGGCGCATGTCTTAAGCCAGCAGAGAAATCCAGC 4997
Qy 673 rAsn-----SerGlnGlnLysSerLeuAlaSer-----GlnSerThrGlnLysG1 688
Db 4998 TAAAGATGTAAGCCCAAGACAGATGAGTACTATCTCTCCCAAGTGAAGCAGATC 5057
Qy 688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGlu-----SerProHI 702
Db 5058 TGGCTTTGACCATTTATCTCTGCACAGTGGAGTGCACCCACTGTTCCCTGAGCCTAC 5117
Qy 702 sProGlnAsnPheGlnSerThrGlnGlnGln-----ThrHisLeuAr 717
Db 5118 TCAGAAATCTGCAGAGTCCATTGGCAAGCAGGTTGAGTGAGAACGAAAGCATGTGAA 5177
Qy 717 gMetGluGlu-----MetVa 722
Db 5178 GAAATCGGAGAACCTTCTCCCATTAAGTACTACCAACAGAGAACCTTCTACACAGT 5237
Qy 722 lThrIleAlaLysSerProLeuPheSerHisAspAspGlnTyr-----IleAl 740
Db 5238 CAGCAACAGAGTCTACAGCATTTCCACACGACATCAGAGTGAAGTAAACGCTC 5297
Qy 740 aGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPh 760
Db 5298 AGATCACCATCAAGCATCAGATTCTAAAGAC----- 5331
Qy 760 eLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGly11 780
Db 5332 ----GTACAGCAGCTCAGATTAGTAAGAGAAATCAGGAGCCCA----- 5373
Qy 780 eGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerAr 800
Db 5374 -TCACCCATCATCTCACCAGCTCAGGAGCAGAGATTCTCTGACAACTCAGAGAGGCT 5432
Qy 800 gGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAl 820
Db 5433 GAGCCCTCTTTT---CCACTGGAGCT---GCACAGAA----- 5466
Qy 820 aArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLysLeuCysAspArgAs 840
Db 5467 -----TCTAGATGATCAGTCCACTGCGCAGTTTCTGACGACCAACAGAGAG 5513
Qy 840 nAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMe 860
Db 5514 TGCTTCATCTCTG-----GAGTGGGAACCTGAGCCACACCTCTATCTG-- 5556
Qy 860 tMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMe 880
Db 5557 -----TCAAAGAGTTTAAAGACATTAATGTTCTGCGCATCTACTACGAAAAGC-- 5607
Qy 880 tGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerTh 900
Db 5608 -----CATCTCCAAAGAGTCCAGGAGCGCATTTTCTGAAGACACTTCTAT 5654
Qy 900 rSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHi 920
Db 5655 T-----GACAATGCCCTGAGTCGACTGACCCCTTGGGAATGAATCTCTGTCAACAA 5705
Qy 920 sGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySe 940

Db 5706 TGGGTACAGTCGAGAGTTCAGATCTTTTCTGAACTCCCTCCTGTGAT-----GGAAA 5759
Qy 940 rLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLy 960
Db 5760 T-----GAAAAGTTGGCTTATCGACGGGACACAAAACAGTCCCG 5801
Qy 960 sLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAl 980
Db 5802 GTCTGCAATATCTATA-----TACAGACC 5825
Qy 980 aLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSe 1000
Db 5826 TATCGACTATGGATCTTTGGGAAAGAACACACAGTGTAGCTTCTTAGAAAAATGTAAGAG 5885
Qy 1000 rSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysG1 1020
Db 5886 GTCA-----CTTACACAAGGAGATATGGAACCAAGTTTCTTAAAGAACCTCG 5936
Qy 1020 yLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAs 1040
Db 5937 C-----TTCTTAAAGATGATTG-----AGGAA 5960
Qy 1040 nPro 1041
Db 5961 CCT 5964

RESULT 12
US-09-945-917-51
; Sequence 51, Application US/09945917
; Publication No. US20030042381A1
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/945,917
; CURRENT FILING DATE: 1998-09-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 4584
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-945-917-51

Alignment Scores:
Pred. No.: 1.63e-09 Length: 4584
Score: 214.00 Matches: 243
Percent Similarity: 34.68% Conservative: 170
Best Local Similarity: 20.40% Mismatches: 414
Query Match: 3.87% Indels: 365
DB: 10 Gaps: 56

US-09-721-114-2 (1-1057) x US-09-945-917-51 (1-4584)

Qy 14 ValGlyThrAsnCysMet-----LeuAlaArgGlyGlyThrGlyAla 27
Db 247 CTGGGTCTCGACTGCTCGAAACTCACCACAAACCGATATCGACAGCGGAAACTTGGGTGCA 306
Qy 28 ValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaGluAlaGly 47
Db 307 GTTCTCCAGCTGCTCTTCTGCTCTCCACTACAGCAG----- 345
Qy 48 ValAspGluProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeu 67
Db 345 ----- 345
Qy 68 LeuGlnLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLys 87
Db 346 -----AAGCTTCGCAACTGAAAAAAGATCAGAGAAA 378

QY 88 CysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgTrpAsp 107
Db : : : : :
379 TTGGAGCACTACCCACATCATATTATGCCACCGCGTCTTAAATACCTCCACAGT 438
QY 108 CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeu 127
Db : : : : :
439 GTCCCAAGCTCAGCAACCGCTTCAGCAACTAACCAAAATCCACATTCACCAAAATGCA 498
QY 128 ProAlaLys---GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr 146
Db : : : : :
499 ACATCCAGGCTTCAGACTCCACAGTCAAGATATCGAAATTTGATTCATCAAGATTTGGT 558
QY 147 PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGly 166
Db : : : : :
559 ATCAAGCCAAAGAGCTGTGACTTAAACCAACCTTCATCATCAACCACTTCATCAAAATAT 618
QY 167 LysAsnAlaAspArgSerThrLeuProLysSerValGlnGlnGlyAsnAsp----- 183
Db : : : : :
619 ACAAAATTCATTCCTG-----CGTGGAGCGTTCGAGTGGCAATTAATATTTGGC 669
QY 184 -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAla 198
Db : : : : :
670 TCGACGATATCCACATCTGCGAAGAGCTTAGAATCATCATCAAGCTACAGCTCATTTTCG 729
QY 199 AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla 218
Db : : : : :
730 AATCTAAACCGACTACTCTCCAACTCCAAACCTTCTAGACCACCAACCCAGCTAGTT 789
QY 219 AsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIle 238
Db : : : : :
790 CGTGTCT-----ACAACTACAAAAATCGAAGC----- 819
QY 239 ThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGlu 258
Db : : : : :
820 -----TCAAGACTAGCGCTCCGAAGCCGTCGAGACCCCAAACTTGTCTCT 867
QY 259 ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu 275
Db : : : : :
868 GTGAGACTATTGAGCAAAACAGAGCCGATTAACAGCGTGTGTGTGTGTGTGTGTGTGTGT 927
QY 276 ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer 290
Db : : : : :
928 CTGAAATTAAGTTATTTCAGTAGCAAAACCCATCTTCCTCATCGAATAGCCCAACCT 987
QY 291 GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProVal 310
Db : : : : :
988 ACGAAGAGCGCGCGCTGCTCAACACAAACTTTGTGAAATTCGTCGCCCAAGT 1047
QY 311 ---SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLys 329
Db : : : : :
1048 AAAAGTGCCTGAAG-----CGCCGACCGAGTAAGCTCGGAAGT 1086
QY 330 ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
Db : : : : :
1087 GCCAGCTATGTGCAAGCTTTGTACGCCAAAGTTTCTACCGTAAACCGAGCGCCCA 1146
QY 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
Db : : : : :
1147 ATCATATCTCAACAA-----GACTCG 1167
QY 362 LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleIleAla 379
Db : : : : :
1168 AAACGATCTCAAGAGCAGTGAAGAGAGTCCGATACGCTGATTCACAGCAGCTC- 1226
QY 379 snAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAlaAsp 398
Db : : : : :
1227 --GCCAAGCTCATCATCAAGGAGTTCCCTAGCATGATTCACATCTTCCAGAGT 1284
QY 399 ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle--- 417
Db : : : : :
1285 TCAAGCTCAGACGAAAGTCTCCGTCATCAGACGATCTTACTCTTAACGCTCCATCGTG 1344

QY 418 -----ProValSerAsnHisThrValGlyGluAsp 427
Db 1345 ACAGCTATCAGACAGCGGATAGCCGCAACACCGGTTTCTCCAAATATTATCAACAGCT 1404
QY 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db 1405 GTTGAGAAAAACCAACACTGCGAGTGAAGAGTGAAGAGCACACCGCAAAAGATGCCA 1464
QY 441 SerAspValAlaAspAspGlySerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db 1465 CTTCCAGCTGTTCCGCCACCGTACACC-----CAGCCCAACA 1500
QY 461 ThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db 1501 ATCGAGTTGTTAGTCCAAATTATGACAT-----AAGAAGTTG 1539
QY 481 Thr-----ProThrAlaSerThrGlnHisAspGluAsnAspThrGluAsnGlyLeu 498
Db 1540 ACAAATGACCCCGTGATATCTGAAAAACAGACCTGAAAGCTCAATCAATGACATC 1599
QY 499 AspThrAsnMetHisLysThrAspVal---CysGlnHisValSerGluIleSerThrGln- 517
Db 1600 GACACG-----ACGAGCTTCCACGCTTCCACCTCTAAAAATCAGTTGTTCCA 1647
QY 518 -----ArgCysSerSerLysGlyLysThzA 526
Db 1648 CTTAAATGACTTCAATCCGACCAACCAACGCTAGCTGTTCTTCTAAACAAGG--- 1703
QY 526 laGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysLysGlyGlyLysLys 546
Db 1704 -----AAAAATCATATGCTGCTCAAG--TCGTTTGGATATGACAGT 1744
QY 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db 1745 CG-----TCGCGCTCTG 1756
QY 566 luAsnSerValLeuSerHis---SerAlaLysValSerPro-----AlaG 580
Db 1757 AAGACTCCATTGTGGTCTCATGCTCGCTCAGCTGACTCGCGCAGCAAAAACCTTCTGTA 1816
QY 580 luHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 600
Db 1817 ATCATTCGCTG-----GAGAGAAGATGGGAAAGATTAAGAT 1855
QY 600 ysGlnLysLeuGluValThr-----ArgGluL 609
Db 1856 CAGAATCCAGCGCTACACTCTGACGCGGTGTTGCGATGTGCGCAAAATCAGGAGGA 1915
QY 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db 1916 AGCTGAAGAATAACGATGACATGCTGCT-----CGAGCACAGACGCT 1960
QY 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db 1961 ATCTGCAACTTCGAAGACAGTCTCTCTGCTGCTGGAATATTCGATAACACAGC 2020
QY 645 leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
Db 2021 TCGACACATATCCACGACGATTTGTCGGAGTACATGCGCAACAGCTGCTCCAAAC 2080
QY 662 spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
Db 2081 AT---ACGACTATTCCACTTTGTCGCTCCATCCAGCTCTTCTCTCAAGCCCGAG 2137
QY 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThr 697
Db 2138 TCCCGAGTCCGCTCCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2184
QY 697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnGlnGlnGln 717
Db 2185 --CAGGAGATGTGTAACAACTTCTGTCGAGTCCGACGACGACGACGACGACGACGACG 2235
QY 717 rgMetGluGluMetValThrIleAlaAlaSerProLeuPheSerHisHis----- 734

184 QY -----SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAla 198
Db TCGACGATATCCATCATCTCGAAGAGTTAGATCATCATCAAGTACAGCTTATTTCG 729
199 QY AsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla 218
Db ATCTAAACCGACCTACTCCCACTCAAAACCTTCTAGACCAACACCCAGCTAGTT 789
219 QY AsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIle 238
Db CGTGTTCGT-----ACAACTACAAAATCGAAGC----- 819
239 QY ThrTrpHisIleGluValAsnGlyValaAspGlnProSerThrProLysLeuSerGlu 258
Db -----TCAAGCTAGCGCTCGGAAGCCGTGAGCACCCCAAACTTGTCTTCT 867
259 QY ValVal-----LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeu 275
Db GTGAAGACTATTGAGCAAAACAGAGCCCGATACAGCGGTGTGTGTGTGGAATG 927
276 QY ValAlaGluGlnCysAsnLeuThrLysAspPro-----AsnProMetSer 290
Db CTGAATTAAGTATTTCAGTAGCAAAAACCCATCTTCTCTCATGATAGCCCAACACT 987
291 QY GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProVal 310
Db ACGAAGAGCGCGCGCTGCTCAACACAACTTTGTGGAATTCGTGCCCGCAGTG 1047
311 QY SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLys 329
Db AAAAGTGGCTGAAG-----CCGCGCACCACTAAGCTGGGAAGT 1086
330 QY ArgSerSerLysSerLys-----ArgLysThrAspLysLys 341
Db GCACGCTATATGTCGAGCTTTGTACGCCAAAGATTTCCTACCGTAAACGGAGCGCCCA 1146
342 QY LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
Db ATCATATCTCAACAA-----GACTCG 1167
362 QY LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleAla 379
Db AAGCATGCTCAAGAGCATGAGAAAGAGTCCGGATACGCTGGATTCAACAGCAGTC- 1226
379 QY snAlaAsnGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAlaAsp 398
Db --GCCAAGCTCATCATCGACGGAAGTTCCCTAAGCATGCTTCCACATCTTCCAAGAT 1284
399 QY ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle---- 417
Db TCAACGTCAGACGAAAAGTCTCGTTCATCAGACGATCTTACTCTTAAACGCTCCATCGTG 1344
418 QY -----ProValSerAsnHisThrValGlyGluAsp 427
Db ACAGCTATCAGACGCGATAGCGGCAACACCGTTTCTCAAAATATTATCAACAGCTT 1404
428 QY -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db GTTCAGGAAAAACCAACTGGCAGTCAAGAGGAGTGAAGAGCACGCAAAAAGATCCA 1464
441 QY SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db CTCCAGCTGTCGCGCACGTGACCC-----CGCCCAACA 1500
461 QY ThrGlySerValHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db ATCGGAGTTGTTAGTCCAAATTATGGACAT-----AAGAGTTG 1539
481 QY -----ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeu 498
Db ACAAATGACCCCTGTATCTGAAAACCAAGAACTGAAAGCTCCAAATCAATGAGCATC 1599
812 QY IuArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832

499 QY AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluIleSerThrGln- 517
Db GACACG-----ACGGACGTTCCACCGCTTCCACCTCTAAATCAGTTGTTCOA 1647
518 QY -----ArgCysSerSerLysGlyLysThrA 526
Db CTTAAATGACTTCAATCCGACCAACACCAACAGTACGATGTTCTTCTAAACAGG- 1703
526 QY laGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGlySer 546
Db -----AAAAATCAGCTGCTCAAG--TCGTTTGGATATGACAGT 1744
546 QY hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db CG-----TCGCGCTCG 1756
566 QY LuAsnSerValLeuSerHis--SerAlaLysValSerPro-----AlaG 580
Db AGACTCCATTTGGCTCATGCTGGCTCAGTGACTCCGCGCAAAAACCTTCTGGA 1816
580 QY LuHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysL 600
Db ATCATTTGCTG-----GAGAGAAGGATGGGAAGAATAAGACAT 1855
600 QY ysGlnLysLeuGluValThr-----ArgGluL 609
Db CAGATCCAGCGCTACACTCTGACCGCGGTGTCGATGTCGCCCAAAATGAGGAGA 1915
609 QY ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db AGCTGAAAGATACATGACATGACTCGT-----CGAGCACAGAACGGCT 1960
629 QY is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db ATCTGACAACTTCGAAGACAGTTCTCTCTGCTGCTGGAATATCCGTAACACAGAC 2020
645 QY leGlnSerLysThrThrAlaAsp-----AspAspCysValIleValAlaAlaLysA 662
Db TCGACGACATATCCAGGACAGATTTGTCGAGTAGACATGGAACACAGTCGCTCCAAC 2080
662 QY spGlySerAspTyrAlaSerSerVal---PheAspThrAsnSerGlnGlnLys----- 678
Db AGGACTATTTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2081
679 QY SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrT 697
Db TCCCCAGTCGCTCCACATCAGTCGATTCGATCTCGAGCAGAA----- 2184
697 QY hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuA 717
Db -----CAGGGAATGTGACAAACTTCTGTCGAGTCCGAAACGAGCCACGTCGSC- 2235
717 QY MetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
Db -----GCCGCTGCCACTTCACTTCGACCAACATTCGCTAA 2272
735 QY AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysL 754
Db GATCCCGGGATCTATCTTCTTCTCA-----CACTATCA---GTGTCAGCTGATA 2323
754 QY ysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC 772
Db AGGACACATGCTATGACACTCAGACACTAGTCGAGACCTTCTTCAAAAACCAAGCT 2383
772 QY ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db ATTCAGGCCAATTT-----CATTCACTTGAT---CGTAAATGCCACTTCAAG 2428
792 QY LysSerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
Db AGTTCATCTCCAGGACAGAGATGGCGGTCTCTTGAGCCCG----- 2472
812 QY IuArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832

Qy 362 LysLeuCys-ArgArgLysProLysLys-ValArgLeu-----LeuSerGluIleAla 379
Db 1168 AAACGATGCTCAAGACGAGTGAAGAGAGTCCGGATACGCTGATTCACACACGTC- 1226
Qy 379 snAlaLeuGlnValGluAspSerArg-SerAspGluValHisArgGluAsnAlaAlaAsp 398
Db 1227 --GCCAACGTCATCATCGACGGAAGTTCCTTAAGCATGATTCACATCTTCCAAGAGT 1284
Qy 399 ProCysGluAspAspArgSerThrIleProValProMetGluValSerMetAspIle--- 417
Db 1285 TCACAGTCAGACGGAAGTCTCCGTCATCAGACGATCTTACTCTTAACGCTCCATCGTG 1344
Qy 418 -----ProValSerAsnHisThrValGlyGluAsp 427
Db 1345 ACAGCTATCAGACAGCCGATAGCCGCAACACCGGTTTCTCCAAATATTATCAACAGCCT 1404
Qy 428 -----GlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db 1405 GTTCAGGAAAAACCAACACTGGCGAGTGAAGGAGTGAAGAGCACAGCGAAAAAGATCCA 1464
Qy 441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysArg 460
Db 1465 CCTCCAGCTGTCGCCACGTCAGACC-----CAGCCCAACA 1500
Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysVal 480
Db 1501 ATCGAGTTGTAGTCCAAATTATGCGCAT-----AAGAGTTG 1539
Qy 481 Thr-----ProThrAlaSerThrGlnHisAspGluAsnAspThrGluAsnGlyLeu 498
Db 1540 ACAATGACCCCGTGATCTGAAACACCAACCTGAAAGCTCAATCAATGAGCATC 1599
Qy 499 AspThrAsnMetHisLysThrAspVal--CysGlnHisValSerGluLysSerThrGln- 517
Db 1600 GACACG-----ACGGACGTTCCACCGCTTCCACCTCTAAATCAGTTGTCCA 1647
Qy 518 -----ArgCysSerSerLysGlyLysThrA 526
Db 1648 CTTAATATGACTTCAATCCGACACACCAACGTCAGTGTCTTCTAAACAAAG- 1703
Qy 526 laGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGluSerT 546
Db 1704 -----AAAATCACAATCGCGCTGTCAAG--TCGTTTGGATATGAGCAGT 1744
Qy 546 hrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrG 566
Db 1745 CG-----TCCGCGTCTG 1756
Qy 566 luAsnSerValLeuSerHis--SerAlaLysValSerPro-----AlaG 580
Db 1757 AAGACTCCATTTGGCTCATGCTCGGCTCAGGTGACTCCGCGACAAAAAATCTCTGGTA 1816
Qy 580 luHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysL 600
Db 1817 ATCATTCGCTG-----GAGAGAAGTATGGGAAGATNAGACAT 1855
Qy 600 ysGlnLysLeuGluValThr-----ArgGluL 609
Db 1856 CAGAATCCAGCGGCTACACTCTGACGCGGTGTTGGATGTCGCCAAAATGAGCGAGA 1915
Qy 609 ysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnH 629
Db 1916 AGCTGAAAGATAACGATGACATGACTCGT-----CGAGCACAGAACCGCT 1960
Qy 629 is-----GluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgI 645
Db 1961 ATCTGTGACAACTTCAAGACAGATTCCTCTGCTGCTGGAATATCCGATAACAACGAGC 2020
Qy 645 leGlnSerLysThrAlaAsp-----AspAspCysValIleValAlaLysAla 662
Db 2021 TCGACGACATATCCACGACGATTTGTCCGGAGTAGACATGCGCAACAGTCGCTCCAAAC 2080

Qy 662 spGlySerAspTyrAlaSerSerVal----PheAspThrAsnSerGlnGlnLys----- 678
Db 2081 AT--AGCGACTATTCCACACTTTGTTCGCCATATCCACCGCTTCTTCTCCTCAAGCCCCGAG 2137
Qy 679 -----SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrT 697
Db 2138 TCCCCAGTCGGTCTCCACATCAGTCGATTCGATCTCGACAGAA----- 2184
Qy 697 hrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGlnGlnGlnThrHisLeuA 717
Db 2185 --CAGAGAAATGTACTACAAACTTCTGCCAGTCGCGAACGAGCCCAACCTGGC----- 2235
Qy 717 xgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis----- 734
Db 2236 -----GCGGTGCGCACTCAACCTTCGGACAACATTCGCTAA 2272
Qy 735 --AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysL 754
Db 2273 GATCCCGGATCATCTATCTTCCA-----CACTTATCA--GFGTCAGCTGATA 2323
Qy 754 ysLeuThrTrpGluGlnPheLysAlaThrThrArg-----AsnSerProAlaAlaThrC 772
Db 2324 AGACACATATGCTATGCACTCAGACATGATGCGAGCTTCTTCAAAAAACCAAGCT 2383
Qy 772 ysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetG 792
Db 2384 ATTCAGGCCAAATTT-----CATTCCTTCTGAT--CGTAAATGCCACCTTCAAG 2428
Qy 792 lysSerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaG 812
Db 2429 AGTTCACTCCACCGACGACAGATGCGGCTCTCTTGGAGCCG----- 2472
Qy 812 luArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluA 832
Db 2473 -----AGACGGGTGCGCAACTCGATGTCGAAATATGAT 2506
Qy 832 laSerLysLeuCysAspArgArgAsnAlaGly-----GlnValValLeuTyrProLysG 850
Db 2507 CTTTCAGATCTTACTCGGCGCTTCCGAGGTGGAGCTCTACTGTATCTATGGAGG- 2565
Qy 850 luSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheP 870
Db 2566 -----ACGTTCCAACTGCACAGACTATCCGATGAAAAATCCCGCCGACAT- 2610
Qy 870 roAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrA 890
Db 2611 -----TCTGCCAAAAGTGAATGGATCCCACTA----- 2640
Qy 890 laHisAsnGlnTyrLysGlySerThrSerThrSerThrSerThrSerThrSerThrSerThr 910
Db 2641 -----TCATCTGGCTAGCACGACGATATGATCT--CTCAATGAGAAGT 2683
Qy 910 leProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu- 929
Db 2684 AC-----GAACATGCTATTTCGGACATGCGACGCTGCTGG 2719
Qy 930 -----ArgProHisProArgValGlyValLeuG 939
Db 2720 AGTGTTCAGAACACATGTCGACTCTAACCAAGAAACAGGAGAACTATGAGCATTTGT 2779
Qy 939 lysSerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSer---- 957
Db 2780 TTGATCTTTTGGACAAAGCTTAGAAAACTCACTCAACACATTCATCGATCCAACTTGA 2839
Qy 958 -----GlyTyrLysLeuGlyValSerThrG 966
Db 2840 AGCTGGAAGGCGAATACGATTCAGCAGGACATTCCTCATTTGAGGGATATTAGCAATC 2899
Qy 966 lyIleThrSerHisGlnMetAsnArgLysGlu----- 976
Db 2900 ATCTTCATCCAACTCAGCTCATGCTAAACGAGGCGCTGCTGAGCTTCTTCGTCAACCAT 2959
Qy 977 -----HisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysT 990

```
Db      2960 CTCTGGATCATGTTGCATCCCATCGATCATCGATCGTCTCGTGAAGCAGCAAGC 3019
Qy      990 rpAsnAlaLeuGlnLeuGlySerValSerSerAlaAspPheLeuSerAlaArgAsnS 1010
Db      3020 AGGAGAGATCAGCTTGAGCTGCTGTTGGCAGAAC----- 3054
Qy      1010 erIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheV 1030
Db      3055 -----AAGAGAGAGCTGGATCGCTCTCTCA-----CTCTCCAAGTTCA 3091
Qy      1030 alArgGlnAspIleCysIleThrAsnLysAsn-----ProAlaAspPheThrThrIleS 1048
Db      3092 CCAAGAG-----AAGACAAGAACTACGACGAGAGCACATATGCCATCAATTI 3139
Qy      1048 erAsnAspAsnGluTyrMetAsp 1055
Db      3140 CCGATCTCAAGGAACCTCTTGAC 3162
```

Search completed: May 17, 2004, 13:07:08
Job time : 2781 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: May 17, 2004, 04:35:29 ; Search time 6194 Seconds
(without alignments)
5095.952 Million cell updates/sec

Title: US-09-721-114-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTCML.....NRKPADFTTISNDNEYMDYR 1057
Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPT.spool/US09721114/runat_14052004_092252_4125/app query.fasta_1.1223
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09721114 @CGN 1.1 4565 @runat_14052004_092252_4125 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estma:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_est1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pia:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_dig:
27: em_gss_vrt:
28: gb_gss1:

29: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	820	14.8	1025	28	BZ795917	BZ795917 PUFHM45TD
C 2	814.5	14.7	888	28	CG355521	CG355521 PUFPL93TB
C 3	716.5	13.0	885	29	CG091207	CG091207 PUFK24TD
C 4	606.5	11.0	787	29	CG345663	CG345663 OGWLW77TH
C 5	567.5	10.3	999	28	BZ795914	BZ795914 PUFHM45TB
C 6	518.5	9.4	766	29	CC603916	CC603916 OGWM55TV
C 7	491.5	8.9	641	28	BZ411511	BZ411511 OGACC43TM
C 8	465.5	8.4	849	28	BZ411504	BZ411504 OGACC43TC
C 9	414	7.5	523	12	BJ246359	BJ246359 BJ246359
C 10	409	7.4	469	29	AG213697	AG213697 Oryza sat
C 11	401.5	7.3	556	13	BU974305	BU974305 HB27110F
C 12	389	7.0	593	13	BQ244810	BQ244810 TABL5035B
C 13	371.5	6.7	653	28	BZ311518	BZ311518 ic68h08.b
C 14	349	6.3	822	29	CG365016	CG365016 OGHAD70TV
C 15	330	6.0	995	29	CG451574	CG451574 OGHAD34TV
C 16	320	5.8	469	9	AL818058	AL818058 AL818058
C 17	310	5.6	734	28	BZ326387	BZ326387 ic68h08.g
C 18	302.5	5.5	753	28	CC160109	CC160109 ig05g11.b
C 19	289	5.2	792	28	BZ422300	BZ422300 id5a111.b
C 20	287	5.2	491	29	AG209720	AG209720 Oryza sat
C 21	286	5.2	905	29	CG365008	CG365008 OGHAD70TH
C 22	275	5.0	970	29	CG032005	CG032005 PUF7A25TD
C 23	270.5	4.9	782	29	CC824592	CC824592 ZMMBB016
C 24	266.5	4.8	903	28	BZ704364	BZ704364 PUFLE78TD
C 25	259.5	4.7	635	28	BH880296	BH880296 ht52d11.b
C 26	251.5	4.6	688	28	BZ776889	BZ776889 ii04h10.g
C 27	245.5	4.4	600	29	CC603904	CC603904 OGWM55TH
C 28	241.5	4.4	703	28	BZ776888	BZ776888 ii04h10.b
C 29	241	4.4	870	29	CG091452	CG091452 PUF8F81TB
C 30	240	4.3	701	28	BZ374744	BZ374744 ie31d12.g
C 31	228.5	4.1	590	28	CC160110	CC160110 ig05g11.g
C 32	219	4.0	5400	29	AY418797	AY418797 Homo sapi
C 33	215	3.9	842	28	BZ705032	BZ705032 PUFCS90TD
C 34	214.5	3.9	587	28	BZ359680	BZ359680 id80b06.b
C 35	214	3.9	628	29	CC744988	CC744988 ZMMBB012
C 36	203	3.7	2030	29	CG756375	CG756375 P051-3-F0
C 37	202	3.7	1879	12	BG168882	BG168882 602320215
C 38	200	3.6	664	13	BQ242025	BQ242025 TABL5035B
C 39	200	3.6	1552	12	BI822914	BI822914 603040353
C 40	196.5	3.6	562	28	BZ652753	BZ652753 OGAM082TC
C 41	195	3.5	4773	29	AY411030	AY411030 Homo sapi
C 42	194.5	3.5	793	28	BZ423675	BZ423675 id52a11.g
C 43	194	3.5	1894	12	BG113113	BG113113 602283772
C 44	193.5	3.5	183	12	BJ252288	BJ252288 BJ252288
C 45	193	3.5	3877	29	AY414473	AY414473 Mus muscu

ALIGNMENTS

RESULT 1
BZ795917/c
LOCUS BZ795917
DEFINITION PUFHM45TD ZM 0.6_1.0_KB Zea mays genomic clone ZMMBTa220H17,
1025 bp DNA linear GSS 17-MAR-2003
ACCESSION BZ795917
VERSION BZ795917.1
KEYWORDS genomic survey sequence.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
Bennetzen, J.

TITLE
JOURNAL
COMMENT
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFH45TB
Contact: Cathy Whitelaw

TIGR Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source

Location/Qualifiers
1..1025
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM061.0.KB"
/clone_lib="ZM 0.6-1.0.KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 5,986-56 Length: 1025
Score: 820.00 Matches: 179
Percent Similarity: 64.60% Conservative: 40
Best Local Similarity: 52.80% Mismatches: 94
Query Match: 14.84% Indels: 26
DB: 28 Gaps: 9

US-09-721-114-2 (1-1057) x BZ795917 (1-1025)

Qy 678 LysSerLeuAlaSerGlnSerThrGln---LysGluLeuGlnGlyHis---LeuAlaLeu 695
Db 982 AAATGTTTCAGCCTCCAGCAATCAAAATGCAAAATGCGCTTCAGGTTCATGTGTACCCATC 923
Qy 696 ThrThrGlnGluSerProHisProGlnAsn-----PheGlnSer 708
Db 922 AAACGGTCTTCATGTTGATCCCTCAAAATTAACCTGCTTCATGACATTTGGAGTGT 863
Qy 709 ThrGlnGlnGlnThrHisLeuArgMetGluGluMetValThrLeuAlaAlaSerSer 728
Db 862 ACCAAGAACACCCAGACACCAATTCACAGGGATGGGGTCCACCAATTCATGCTACCTCA 803
Qy 729 ProLeuPheSerHisAspAspGlnTyrlleAlaGluAlaProThrGluHisTTPGly 748
Db 802 CCTATGTTTCACATCAT-----CAACATATGCTGAAGTGTCTGCTCGAGCTGGAGG 749
Qy 749 ArgLysAspAlaLysLysLeuThrTTPGluGlnPheLysAlaThrThrArgAsnSerPro 768
Db 748 AATAACGGGGAAGAGTTAATGTGGGATTCCTTCAAGACAGCTTCACAGAAATTCACCA 699
Qy 769 AlaAlaThrCysGlyAlaGlnPheArgProGlylleGlnAlaValAspLeuThrSerThr 788
Db 688 ACATCGCATATGTTTTCATTCGGAACAGGCTTCACAGAGTTCATTCCTCCAGT 629
Qy 789 HisValMetGlySerSerSerAsnTyrlleAlaSerArgGlnProValleAlaProLeuAsp 808
Db 628 CATGCTTATGAGCTTCTAGTAACATATGAGCTCACCAGCCAGTAATTTGGTGGCTAGAT 569
Qy 809 ArgTyrlleAlaGluArgAlaValAsnGlnValHisAlaArgAsnProSerThrleAla 828
Db 568 CAG-----CCGAGAGTGTTCACAGCAACATCG 539
Qy 829 ThrMetGluAlaSerLysLeuCysAspArgArgAlaGlnValValLeuTyrlPro 848
Db 538 ACTATGAGGCTGTAGTGTATGATCGAGGATTCCTGACAGCTCAGGCTTACCCA 479

Qy 849 LysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer 868
Db 478 AAGAGCCCTATGCCCTGCAACACATTTCTGAGATTGATGGACTCATCAACACTCCAGAC 419
Qy 869 PheProAsnTyrlleGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHis-----Asn 886
Db 418 TTCACAAACTATCAAGAGCTAACAGAGCCAGAGTGAACATCAACACACAAACTCTGGAT 359
Qy 887 SerGlnTyrlleAla-----HisAsnGlnTyrlleGlySerThrSerThrTyrlleGlySerAsn 905
Db 358 TCACATATACACAGCATGATCATGATATATGATGCTCCTGAGCAGCATCATATGAGACAC 299
Qy 906 LeuAsnGlyLysleProLeuThrPheGluAspLeuSerArgHisGlnLeu---HisAsp 924
Db 298 ATAACTGGAAGAGTTCCACCGACACTGCAAGACTTATCTCGGCGTCAGGTCCAGCAAAAC 239
Qy 925 LeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuGlnLys 944
Db 238 CTGCACAGGCTTTAGCCCTCATCTCGGGTGGGCGTCTTGGTTCGTTGCTGAGCAG 179
Qy 945 GluileAlaAsnTrpSerGluAsnCysGly---ThrGlnSerGlyTyrlleGlyVal 963
Db 178 GATATCGCAAACTGTTATGAAACTCGGGCGCCACAGCTCTGGATACAGGCTGGCGTT 119
Qy 964 SerThrGlyLeThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSer 983
Db 118 FCTAAGGAGCGGCATCGCTTGATATGACACACATGGAACATACGAGACCTTGAACTCA 59
Qy 984 GlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerAla 1002
Db 58 GGAGTGTTCACAGCAGGATGGATGCGCTTCAATGGGTTCTGTAGCTCTCTTGGC 2

RESULT 2

LOCUS

CC355521

DEFINITION

genomic survey sequence.

ACCESSION

CC355521.1

VERSION

GI:30824921

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 888)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and

Bennetzen, J.

Maize Genomics Consortium

Unpublished (2003)

Other_GSSs: PUFH45TB

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..888

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZM061.0.KB"

/clone_lib="ZM 0.6-1.0.KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN

Alignment Scores:

genomic survey sequence.

CG091207 1 G1:33973501
GSS.
Zea mays
Zea mays
Zea mays

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 885)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUJFK247E
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
Location/Qualifiers
1..885
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB7a0656C23"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 1,18e-47 Length: 885
Score: 716.50 Matches: 160
Percent Similarity: 60.97% Conservative: 29
Best Local Similarity: 51.61% Mismatches: 88
Query Match: 12.97% Indels: 33
DS: 29 Gaps: 9

US-09-721-114-2 (1-1057) x CG091207 (1-885)

QY 753 LysAlaSerLeuThrTrpGluGlnPheLysAlaThrThrArgAsn-----SerProAlaAla 770
Db 27 AAGAAGCTATGTCGGGATCTTTTCAGACAGCTTCAAGGATCCACCATCCACCACCG 86
QY 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 87 TCATATGCTTTCAATTTCGGAACAGGCTTCAAGAGTTTGATTTCAGCTCCCATTCATGCA 146
QY 791 MetGlySerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyr 810
Db 147 -----TATGGAGCTCACCGCCAGTAAATTGTGGCCGTAGATCAG--- 185
QY 811 AlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMet 830
Db 186 -----CCGAGAAGTGTTCGAAGCACACAGCGTTGGCTATG 218
QY 831 GluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLysGlu 850
Db 219 GAGCGTGTAGAGCTGTACGATCAAAAGATCCCTGCAGCAGCGCCCTTTTACCCAAAGAG 278
QY 851 SerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro 870
Db 279 CCCATGCTCGAACACATCTCTCTGGGCTGATGGACTCATCAACAGCTCCAGGCTTCACA 338
QY 871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu-----SerGlnLeuHisAsnSerGln 885
Db 339 AACTACCAAGAGCTAATAGCGGCAGATGAGCTTCAACACGCAAGCTCTCGGCTCCGAG 398

Pred. No.: 1,135e-55 Length: 888
Score: 814.50 Matches: 169
Percent Similarity: 69.83% Conservative: 37
Best Local Similarity: 57.29% Mismatches: 17
Query Match: 14.74% Indels: 72
DB: 28 Gaps: 6

US-09-721-114-2 (1-1057) : CC355521 (1-888)

QY 767 SerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThr 786
Db 2 TCACCAACATCGTCATATGTTTCAATTTCGGAACAGGCTTCAAGAGTTTCATTCATCT 61
QY 787 SerThrHisValMetGlySerSerAsnTyrAlaSerArgGlnProValIleAlaPro 806
Db 62 CCAGCTCATCTTATGGAGCTTCTATTAATACTATGCAGCTCACAGCCAGTAATTGGCT 121
QY 807 LeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThr 826
Db 122 ATAGATCAG-----CCGAGAAGTGTTCGAAGCAC 151
QY 827 IleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeu 846
Db 152 ACATCGACTATGAGGCTGTAGTGTATGATCGAAGGATTCCTGGACAGTCAGGCTC 211
QY 847 TyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeu 866
Db 212 TACCCAAAGAGCTATGCTCGAACACATCTCTGAGATTGATGACTCATCAACAGCT 271
QY 867 AlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHis--- 885
Db 272 CCAGACTTCACAACTATCAAGAGCTAAACAGCGCCAGATGGAACATCAACACAAACT 331
QY 886 ---AsnSerGlnTyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGly 903
Db 332 CTGATTCACAAATATACACAGCATGATGATGATGAGGTCTACTGAGCATCATATGGA 391
QY 904 SerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeu--- 922
Db 392 AGACACATAACTGGAAGGTTCCACCGACACTGCAAGACTTATCTCGGGCTCAGGTCCAG 451
QY 923 HisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeu 942
Db 452 CAAAACCTGCACAGGCGCTTTACGCCCTCATCTCCGGTGGGGTGGTGGTGGTGGTGG 511
QY 943 GlnLysGluIleAlaAsnTrpSerGluAsnCysGly---ThrGlnSerGlyTyrLysLeu 961
Db 512 CACGAGATATCGCAAACTGGTATGAAACTGGGGCCGCCACAGTCTGATACAGGCTG 571
QY 962 GlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeu 981
Db 572 GGGCTTTCTAAAGGAGCGGATCGCTGTATGATGAACAGACATGGAACATACGAGACCTTG 631
QY 982 AsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer 1001
Db 632 AACTCAGAGTGTTCACAGCAGGATGAATGCGCTTCAATGGGTCTCTGTAGCTCTCTT 691
QY 1002 AlaAsp-----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLys 1019
Db 692 GCGCATCCAGATACCGGTTGCCAAGGTATGTTACAGGTCCAGCTTGTACAGTGGCAAT 751
QY 1020 GlyLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLys 1039
Db 752 GGGAGAACGGTTCACCGCTTGGTAAGCTGTGTGCGAAAGGATATCTGTGTGATACAGA 811
QY 1040 AsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db 812 AACCCAGCTGATTTACTGTAAATTAGTGACATGATGATGATATG 856

RESULT 3
CG091207 885 bp DNA linear GSS 20-AUG-2003
LOCUS PUJFK247D ZM 0.6 1.0 KB Zea mays genomic clone ZMMB7a0656C23.
DEFINITION PUJFK247D ZM 0.6 1.0 KB Zea mays genomic clone ZMMB7a0656C23.

DEC 17 1964

[illegible]

[illegible]

Db 286 ACAAGCGCGCTGATGCGACAGAGGAGGAACTACGAGACCTTGAGC----- 333

Qy 987 SerHialysTrpAsnAlaLeuGlnLeuGlySerValSerSerAla-----AspPhe 1004

Db 334 TCGGAGGATGGAACACCTCGAGTCTGGTTCCTGTAGCTCTGCGCAATCTGGAGTAC 393

Qy 1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024

Db 394 CGGTTTCCATGTTAGTACGGGTACCGCTTCGACCATGGCATGGAGGACCATTCAT 453

Qy 1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPhe 1044

Db 454 CCGTTGATAGTACGTCGTGAGAGGATATCTGTGTGACTAACAGAAACCGCGGATTC 513

Qy 1045 ThrThrIleSerAsnAspAsnGluTyrMet 1054

Db 514 ACTGTAATTAGTGACAGACGAGTACATG 543

RESULT 8

BZ411504/c

LOCUS

DEFINITION

OGACC43TC ZM 0.7 1.5 kb Zea mays genomic clone ZMWBa0027112,

genomic survey sequence.

ACCESSION

BZ411504

VERSION

BZ411504.1 GI:26044995

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Zea mays

REFERENCE

1. (bases 1 to 849)

White, G.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

OTHER_GSSs: OGACC43TM

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1. 849

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZM 0.7 1.5 kb"

/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN

source

Alignment Scores:

Pred. No.: 2,476-27 Length: 849

Score: 465.50 Matches: 106

Percent Similarity: 64.40% Conservative: 17

Best Local Similarity: 55.50% Mismatches: 52

Query Match: 8.42% Indels: 16

DB: 28 Caps: 6

US-09-721-114-2 (1-1057) x BZ411504 (1-849)

Qy 871 AsnTyrGlyThrSerSerArg-AsnGlnMet-----GluSerGlnLeuHisAsnSerG1 888

Db 845 AACTACCAAGAGCTATAGGGCGGAGATGACCTTCAACCGCAGCTCTCGGCTCGCA 786

Qy 888 nTyrAla---HisAsnGlnTyrLysGlySerThrThrSerThrSerTyrGlySerAsnLeuAs 907

Db 785 GTACACAGCATGACATTCACATCGGTACCGAGCACGTATATATGGAAGCAA----- 731

Qy 907 nGlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHi 926

Db 730 -----CGCTCAGACTGGAAGACTGTCTCGGCGTCAATTCAGCAAGACTTGGC 681

Qy 926 sarGProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGlu11 946

Db 680 CAGGCTTTACGCTTCACCTCGTCTCGGTGTCGCGTTCATTCTGTCAGCAGAGAT 621

Qy 946 eAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrG1 966

Db 620 CGCAAACTGGTCTGGGAAGCTGCGGCGCGAGCTGCGGTACAGACTAGGTGATGTAAGG 561

Qy 966 yIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPh 986

Db 560 GACAGGCGCGCTGATGCGCAACAGACAGCAAACTACGAGACCTTGAGC----- 512

Qy 986 eSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla-----AspPh 1004

Db 511 -TCGGCAGGATGGAACACCTCGAGTCTGGTTCCTGTAGCTCTGCTGCCAATCTGGAGTA 453

Qy 1004 eLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHi 1024

Db 452 CCGGTTTCCATGTTAGTACGGGTACCGCTTCGACCATGGCAATGGAGGACCATTC 393

Qy 1024 sProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPh 1044

Db 392 TCGTTGATAGTACGTCGTGAGAGGATATCTGTGTGACTAACAGAAACCGCGGATTT 333

Qy 1044 eThrThrIleSerAsnAspAsnGluTyrMet 1054

Db 332 CACTGTAATTAGTGACAGACGAGTACATG 302

RESULT 9

BZ246359

LOCUS

DEFINITION

BZ246359 Y. Ogiwara unpublished cDNA library, Wh_f Triticum aestivum cDNA clone whf22j14 5', mRNA sequence.

ACCESSION

BZ246359

VERSION

BZ246359.1 GI:20058037

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

REFERENCE

1. (bases 1 to 523)

Ogiwara, Y. and Murai, K.

Expressed genes in Triticum aestivum

Unpublished (2002)

Contact: Tadasi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tschin@genes.nig.ac.jp.

Location/Qualifiers

1. 523

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="whf22j14"

/tissue_type="spike at flowering date"

/dev_stage="Peekes' scale 10.5.1"

/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"

ORIGIN

source

Alignment Scores:

Pred. No.: 1,91e-23 Length: 523

Score: 414.00 Matches: 85
Percent Similarity: 55.16% Conservativity: 16
Best Local Similarity: 54.84% Mismatches: 54
Query Match: 7.49% Indels: 0
DB: 12 Gaps: 0

US-09-721-114-2 (1-1057) x BJ246359 (1-523)

QY 708 SerThrGlnGluGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSer 727
DB 51 TCTRCCAGGAGACACAGGACATCGGGGATGGAGATATGTCATATGTCGACCC 110
QY 728 SerProLeuPheSerHisAspGlnThrIleAlaGluAlaProThrGluHisTTP 747
DB 111 TCACCACTATTATCACCAGACAGACAGCAGCTCTGCTGAAGCAACACCTGAGCGGTGG 170
QY 748 GlyArgIysAspAlaIysLeuThrTTPGluGlnPheIysAlaThrThrArgAsnSer 767
DB 171 AGCCATATGGAGACAAAGAGTCGATGTGGAACTTTCAAGGCACCTTCACAGGGATT 230
QY 768 ProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSer 787
DB 231 TCACCAACACACAGGTGTGCTGCTCAGCTCAGACCTAGCATTGACACGGTTGATTAACTTAT 290
QY 788 ThrHisValMetGlySerSerSerAsnThrAlaSerArgGlnProValIleAlaProLeu 807
DB 291 ACTGATGGGGGAGCTAATCGTTATTATCCCACTCGCCAGCAGTAATTTGCACACT 350
QY 808 AspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIle 827
DB 351 GACCACTATCAATAGACAGTAACTCAGTCAGCCAGCAGAGTTTCCAGTTTCAGTG 410
QY 828 AlaThrMetGluAlaSerLysLeuCysAspArgAsnAlaGlyGlnValLeuTyr 847
DB 411 TCACCACTGGAAGCTGGTAATCTGTGTGATGGAGAAATGTTGACATTCAGGTTTAT 470
QY 848 ProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAsp 862
DB 471 CCAGAGAAACCATGCTGCTACTACGCCCCCTGATGATGACCGAG 515

RESULT 10
LOCUS AG213697 469 bp DNA linear GSS 09-AUG-2003
DEFINITION Oryza sativa (japonica cultivar-group) DNA, clone: T10466T, 3' flanking sequence of Tos17 insertion in rice strain ND6047, genomic survey sequence.
ACCESSION AG213697.1 GI:32360887
VERSION AG213697.1
KEYWORDS GSS.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1 Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., Shinzuka, Y., Onosato, K. and Hirochika, H. Target Site Specificity of the Tos17 Retrotransposon Shows a Preference for Insertion within Genes and against Insertion in Retrotransposon-Rich Regions of the Genome Plant Cell 15 (8), 1771-1780 (2003)
JOURNAL MEDLINE 22779046
PUBMED 12897251
REFERENCE 2 (bases 1 to 469)
AUTHORS Miyao, A., Miyazaki, A., Yamashita, Y. and Hirochika, H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Akio Miyao, National Institute of Agrobiological Sciences, Molecular Genetics, 2-1-2, Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: miyao@affrc.go.jp, URL: http://tos.nias.affrc.go.jp/, Tel: 81-298-38-7020, Fax: 81-298-38-7020)
FEATURES Location/Qualifiers
source 1. 469

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="ND6047"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="T10466T"
/genomic DNA"
/note="The 3' end of retrotransposon Tos17 was found immediately upstream of this sequence."
misc_feature 1
/note="This insertion point is base 1732 in AF326768."

ORIGIN

Alignment Scores: 4.17e-23 Length: 469
Pred. No.: 409.00 Matches: 79
Score: 100.00% Conservativity: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 7.40% Gaps: 0
DB: 29
US-09-721-114-2 (1-1057) x AG213697 (1-469)
QY 498 LeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGln 517
DB 1 CTGACACAAATATGATTAAGACAGATGCTGTGACCATGATCAGAAATCTCCACACAG 60
QY 518 ArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAla 537
DB 61 AGGTGCTCATCAAGGGGAAACACAGCGGCTTTGAGTAAAGGGGAAACACATTCAGCTGCT 120
QY 538 SerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAla 557
DB 121 AGTACCAATATGGTGGTGAAGACCCAGAAATGTCAGAACATCATGTACTCAGCGCA 180
QY 558 GluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal 576
DB 181 GAAGATCAATGCCAGATGGAAACCCGAAACTCTGTTCTGAGTCACTCGGCAAGGTA 237

RESULT 11
LOCUS BU974305 556 bp mRNA linear EST 22-OCT-2002
DEFINITION HB27110r BC Hordeum vulgare subsp. vulgare cDNA clone HB27110
S-PRIME mRNA sequence.
ACCESSION BU974305
VERSION BU974305.1 GI:24225098
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
REFERENCE 1 (bases 1 to 556)
AUTHORS Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.
TITLE Barley ESTs from developing seeds
JOURNAL Unpublished (2002)
COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 556 Std Error: 0.00
Plate: 27 row: 1 column: 10
Seq primer: M13rev, Location/Qualifiers
1. 556
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"


```

FEATURES
  source
    Class: sheared ends.
    Location/Qualifiers
      1..822
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone_lib="ZM071.5 KB"
        /notes="Vector: pBCSK-1; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 6,556-18 Length: 822
Score: 349.00 Matches: 109
Percent Similarity: 47.49% Conservative: 33
Best Local Similarity: 36.45% Mismatches: 90
Query Match: 6.32% Indels: 68
DB: 29 Gaps: 15

US-09-721-114-2 (1-1057) x CG365016 (1-822)
Qy 63 GYTYrValAlaLeuGlnLysAspProLysPheCysSerLeuSerArgIlePhe 82
Db 3 GGTATGTTGCTATGCTTCAGAGAGGATCCAAACCTTGCTCT---CCTCATATTTTC 59
Qy 83 HisAspGlnLysCysAspGluHisLysAlaSerSerProPheSerValAlaLys 102
Db 60 CATAGCAGCCCAATATGATGAGCACCAT---GGTTCGTCGCCACTTTGGTGTCAAG 116
Qy 103 PheArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThr 122
Db 117 TATCGTCGATGGATGCTCAAAATGCTTGAAGAGATGAAAGTTTCAGGCCATCGACT 176
Qy 123 AlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPhe 142
Db 177 ACATCTGGAATGTTTCTATCGAGCAGGATGAATGAATGATGGTCTCTATTTCAATT 236
Qy 143 ValArgSerThrPheValProAlaSerValGlySerGlnLysVal---SerProSerThr 161
Db 237 GTTTCGG-----ATTTTGCTAATAGGTGTATCCAGAGAAATGTTTATTCGAACAA 290
Qy 162 GlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGly 181
Db 291 CATCATCTCAGGGAT-----GATCAATGACCTTTCAGAACGCCCAAGATGC 344
Qy 182 AsnAspSerLysCysAsnAlaPro----- 189
Db 345 AAT---TCAAAATGCAGTTCTCTCTGGCAACAAAGCAATTACTGCGATGAATATCCAGTA 401
Qy 190 ----- 192
Db 402 GCTGAAGAAATGTTCTAGAGGCACCTGTTGAAGAGAGTGTCCAGCGACAGAGATATG 461
Qy 193 AsnGlyAlaAlaGlu-----AlaAsnThrAsp 201
Db 462 GCTGGCGATGTTGAATTTCTTAATTAATCTTTCTCTTGTATACAAATAAGCTAATAT-GAC 520
Qy 202 SerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSer 221
Db 521 AGTATTTCTCAGATTTACAGGCTCTCCCAATATATATGATGTTCTCAGCAAAATATCTG 580
Qy 222 GluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrpHis 241
Db 581 ---AATGTGTCTCCCAAGGATCCAGCGATTTTACCTGTGATGTTTCAAACCATATCAACC 637
Qy 242 IleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeu 261
Db 638 ATTGAAGAAATGGCCAAATAATCCATGTAG---CCAAACTTCGT-----GTG 685
Qy 262 LysArgAsnGluAspGlnAsnGlyLysThrGluThrLeuValAlaGluGlnCysAsn 281

```

686 ATGCCAATGAAGATGAGAGT-----AACATAGTTCAGAGTTCCTAAT 730

282 LeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCys 301

731 TTT-----GACCCCAAT-----GAGTCTAATGTTTGTAAACCATTA 766

302 AsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320

767 TCTTGCCCAAAAGGC---AAACAATATCTGGGCACAAAGAGCAGCCAGCTGCAAC 820

RESULT 15

CG451574 995 bp DNA linear GSS 17-SEP-2003

LOCUS CG8AD34TV ZM 0.7_1.5 KB Zea mays genomic clone ZM58MA0848F20,

genomic survey sequence.

CG451574

CG451574.1 GI:34836574

GSS.

Zeasays

Zeasays

Zeasays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 995)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..995

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZM071.5 KB"

/clone_lib="ZM 0.7_1.5 KB"

/note="Vector: pBCSK-1; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ALIGNMENT SCORES

Pred. No.: 2,956-16 Length: 995

Score: 330.00 Matches: 102

Percent Similarity: 46.44% Conservative: 35

Best Local Similarity: 34.58% Mismatches: 72

Query Match: 5.97% Indels: 87

DB: 29 Gaps: 14

US-09-721-114-2 (1-1057) x CG451574 (1-995)

Qy 61 IleArgGlyTyrValAlaLeuLeuGlnLysAspProLysPheCysSerLeuSerArg 80

Db 242 GTCAGGGTATGTTGCTATGCTTCAGAGAGAGTCCAAACTTGC-----TCTCAT 295

Qy 81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100

Db 296 ATTTTCCATAACAGCCCAATATGATGAGCACCAT---GATTCTGCCCACTTTTGGTG 352

Qy 101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120

Db 353 TCGAGTATCATCGATGGGATGCTCAGATGCTCTTGATAGTGTGAGAGTTTCAGGCCAT 412

Qy 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140

Db 413 AGGCCAACATCGAAAATGTTTCTATGACGACGAATGAATGAATGATGGCTGCTCTATT 472
Qy 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal---SerPro 159
Db 473 TCATGTTGCGG-----ATTGTGCTAATAGTGTGTGATTCGAGAGACTGTTTCTTCGC 526
Qy 160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db 527 ACACAGCAATCATCTCAAGGAAT-----GATCGATTGACCCCTTCAAAGACTGCCCAA 580
Qy 180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLys----- 192
Db 581 AAATGCAAT---TCAAAATGCGAGTTCTCTCGCAACAAGCAATTAATCTGCGATGAATGTT 637
Qy 193 -----Asn 193
Db 638 CCACTAGCTGAAGAAATGTTCCAGAGGCACTCGTTGATACGAGGGTTCATCTATAGAA 697
Qy 194 GlyAlaAlaGlu-----AlaAsn 199
Db 698 GGTATGGCTGACATGTTGAATCTTAATTACTTTTGTGTATACAAAGCTAAT 757
Qy 200 ThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsn 219
Db 758 AT-GACAGTATTTCTCAGTTTTCAGGGCTCTCCCAATAGTATGATTGTGTCAGCAAT 816
Qy 220 ValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThr 239
Db 817 ATCTTG---AATGCTGTCTCGAAGAATGTCAGCGATCTACCTGAT----- 858
Qy 240 TrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluVal 259
Db 859 ---GATGTTGAAGAAATGGCACAGATCCACAGTCTA-----AAACCT 903
Qy 260 ValLeuLysArgAsnGlnAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGln 279
Db 904 TGTGTGGCGCCAAATCAAGATTAGAATAACATAGCTAATGATGACCTC----- 951
Qy 280 CysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGlu 299
Db 952 -----ATGAGCCTAAT----- 963
Qy 300 GlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
Db 964 GTATGT-----AAACCAATATCTGGGCACAAA 990

Search completed: May 17, 2004, 09:59:24
Job time : 6235 secs